This project has received funding from the European Union's Horizon EUROPE research and innovation program under grant agreement No. 101086512.

This project was funded by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee [grant number 10066637].

Call: HORIZON-CL6-2022-GOVERNANCE-01
Project 101086512

INN WATER

Promoting social innovation to renew multi-level and cross sector water governance

D4.3: Methodology for analysing the socio-economic performance of Household Water Demand Policy

Michel PAUL (CEMOI – University of Reunion Island)

Delivery date: 08/06/2026

Disclaimer

This document reflects only the author's view. The European Commission is not responsible for any use that may be made of the information it contains.

Intellectual Property Rights

Project Consortium

INN WATER

Document information

Programme	HORIZON Research and Innovation Action - HORIZON-CL6-2022-GOVERNANCE-01-06
Grant Agreement N°	101086512
Project Acronym	**InnWater**
Project full name	Promoting social INNovation to renew multi-level and cross sector WATER governance
Start of the project	1 March 2023
Duration	36 months
Project coordination	Julie Magnier, Office International de l'Eau, OiEau
Deliverable	**D4.3:** Methodology for analysing the socio-economic performnace of household water demand management policy
Work Package	**WP4:** Digital tool for water governance
Task	**Task 4.3:** Domestic water tariff dashboard
Lead Beneficiary	University of Reunion Island
Author(s)	Michel PAUL (CEMOI – University of Reunion Island)
Contributor(s)	Geoffrey JOHNEN (CEMOI – University of Reunion Island) : sections 12.2, 12.3 and 12.4
Quality check	Andras KIS (REKK)
Planned Delivery Date	31/10/2024
Actual Delivery Date	08/06/2026
Citation	PAUL M., (2026): *Methodology for analysing the socio-economic performance of household water demand management policy,* Deliverable D4.3, EU Horizon InnWater Project, Grant agreement No. 101086512
Dissemination Level	**Public**

Revision history

Version	Date	Author(s)/Contributor(s)	Comments
V1	11/07/2023	Natacha Amorsi (OiEau)	
V2	17/10/2024	Michel PAUL (CEMOI – University of Reunion Island)	First draft version
V3	30/06/2025	Michel PAUL (CEMOI – University of Reunion Island)	First Revised version
FV	31/07/2025	Michel PAUL (CEMOI – University of Reunion Island)	Final version
FV2	08/06/2026	Michel PAUL (CEMOI – University of Reunion Island)	Final version considering review comments

Related deliverables

This document is linked to the *D4.2 Modelling cross-sectoral interactions with water at river basin level*, as the macroeconomic simulation level, as well as to the D4.4 and D4.5 *InnWater Governance platform* deliverables, since the tool will be available on this platform. Learning documentation and training exercice arising from D4.3 were also created, and will be briefly introduced in *D6.4 Communication and Dissemination report and monitoring V2*.

INN WATER

TABLE OF CONTENT

EXECUTIVE SUMMARY

This document, Deliverable D4.3: Methodology for analysing the socio-economic performance of Household Water Demand Policy, is part of the InnWater project (Promoting social INNovation to renew multi-level and cross sector WATER governance). It falls under Work Package 4: Digital tool for water governance and Task 4.3: Domestic water tariff dashboard. The project itself is funded by the European Union's Horizon EUROPE research and innovation program and UK Research and Innovation.

The primary objective of Task 4.3 is to develop a microsimulation model (MMS) to assess the socio-economic performance of water pricing policies for domestic uses, particularly the Increasing Block Tariff (IBT) scheme. This particular deliverable sets out the methodology for evaluating a single pricing policy, mostly at a municipal scale, and also some elements related to the evaluation of a set of water pricing policies (with the up-scaling operation of the basic/ disaggregated MMS). Its broader aim is to inform public decision-making on water pricing, providing stakeholders with useful and relevant information on socio-economic consequences for both diagnostic and exploratory purposes.

The methodology involves a digital tool with three main features. It is based on econometric estimates of local household water demand functions (using data from Reunion Island for demonstration purposes) which are crucial for identifying (and quantifying) consumption determinants such as family size, household income, and tariff parameters. It employs academic indicators from water economics and other fields of social sciences to measure performance across five key areas related to the European Water Framework Directive (EU-WFD):

Affordability: assessing whether households can meet their water needs at socially acceptable economic conditions, measured by indicators like the Conventional Affordability Ratio (CAR) and Potential Affordability Ratio (PAR);

Incentive effect of pricing: measuring the tariff's ability to encourage water conservation, including the impact on consumption and the prevalence of overconsumption due to tariff misperception;

Economic efficiency: analyzing the welfare gains and losses generated by IBT implementation, using concepts like aggregate (social) surplus and consumer surplus;

Equity: examining the redistributive impacts of the water pricing policy, including cross-subsidy system (implemented by progressive pricing), social targeting quality, and effects on household income inequalities (e.g., using Lorenz curves and Gini index),

Cost recovery (quality of the funding): évaluating the financial sustainability of the domestic water (and wastewater) service, by analyzing operating results and financing structures.

And, finally, it is based on French regulations governing the pricing of drinking water and wastewater services which includes some basic principles like the "Water pays for Water" principle (with service funding provided by domestic users and sales rather than through taxation) or the "Universality of service" principle (the latter prevents social pricing of water domestic uses with specific tariff for poor household).

The MMS is structured around five modules: Population, Tariff, Demand, Invoice, and Evaluation, each handling specific data and calculations. It is developed as a web application in Python and Angular. The project plans to release the tool under a gpl-3.0 license, promoting open access and collaborative development.

Part of this deliverable focuses on the "scaling up" operation (which, to the best of our knowledge, is a first) . This aims to transform the MMS from a municipal-level tool to one capable of providing information at larger geographical scales, such as river basins. This expansion is motivated by the need for macro-level understanding, better targeting of support measures, and support for multi-level water governance. The scaling-up design involves encapsulating the basic MMS, creating PopulationResults objects and Aggregators to handle diverse aggregation methods (e.g., weighted average, variance decomposition), and integrating a geospatial database (SQLAlchemy) for French administrative divisions (from basin to IRIS INSEE) to deal with crucial spatial dimension. This comprehensive architecture allows for the central execution of municipal simulations and subsequent aggregation to user-defined administrative levels. Once completed, this expansion will enable the measurement of spatial inequalities and aggregate performance across specific sub-areas, utilizing new geospatial indicators.

The EU-added value of this work lies in its potential to significantly improve water governance related to water price setting. By integrating academic methodology (and knowledge) into operational digital tools, it provides relevant information for diagnosing existing policies, exploring proposed ones and highlighting the actual trade-offs between various policy objectives. Once completed, the up-scaled MMS is also expected to (i) enable better targeting of support measures, by identifying spatial disparities and inequalities, and (ii) facilitate multi-level governance and dialogue among diverse stakeholders by providing a common and shared evaluation methodology.

This work is interlinked with other InnWater project actions as it directly feeds into the development of the broader digital tool (WP4). Ultimately, it supports InnWater's overall goal of promoting social innovation in multi-level and cross-sector water governance by providing robust, data-driven decision support. The ultimate goal is to serve as a Decision Support Tool (DST) to inform public decision-making. Some of the information provided by the DST could also be used to induce and foster Citizen Engagement in water grovernance

.

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

A Tariff	“Tarif Assainissement” (Wastewater Tariff)
CA	Consortium agreement
CTM	"Coût Total Moyen" (Unit Total Cost)
CV	"Cout Variable" (Variable Cost)
CVM	"Coût Variable Moyen (Unit Variable Cost)
EC	European Commission
EP Tariff	“Tarif Eau Potable” (Drinking Water Tariff)
EPA Tariff	“Tarif Eau Potable et Assainissement" (Drinking Water and Sanitation Tariff)
EU-WFD	European Water Framework Directive
IBT	Increasing Block Tariff
MMS	“Modèle de Micro-Simulation” (Microsimulation Model)
RV	"Recette Variable" (Variable Revenue)
RVM	"Recette Variable Moyenne" (Unit Variable Revenue)
TBSE	“Tarification Binôme Stucturellement Equilibrée” (Two-part Structurally Balanced Tariff)
WP	Work Package

I – INTRODUCTION

One of the deliverables of the InnWater project is a microsimulation model (MMS) that enables to assess the socio-economic performance of the water pricing policy (Linear Tariff, Two-part Tariff, Increasing Block Tariff) for domestic uses which is implemented by a water company, the stakeholders involved in water price setting more generally, in the area (most often, a city) in charge. This assessment is carried out by means of a dashboard and the use of "appropriate" indicators in each of 5 major fields of analysis related to the European Water Framework Directive (EU-WFD): Affordability; Incentive effect of pricing; Economic efficiency; Equity; Cost recovery (quality of the funding). This report presents the prototype (and the related methodology on which it is based) that was created during the implementation of the project.

This digital tool has three main features. The first is to be based on econometric estimates of local household water demand functions, that is on the estimated causal relationship that links household water consumption to its main determinants (size and composition of the family, level of income, tariff parameters ...). Coupled with the database used for its estimation, knowledge of the water demand functions of local households does provide essential information to carry out some relevant public policy evaluations, whether for diagnostic or simulation purposes. The information in question refers mainly to:

- the measurement of the volumes of water that are necessary to meet the basic needs of households (and that change across the population);
- the extent of overconsumption (linked to tariff misperception when the pricing structure is complex) ;
- the price-responsiveness of water demand (including the degree of the proper perception of the tariff).

The tool then uses, for demonstration purposes, a specific econometric model, namely the one estimated for Réunion Island by Binet, Carlevaro and Paul [2014], referred as BCP in the following.

The second main feature of the tool is to use academic indicators that are commonly applied in the field of water economics or in other areas of economics (like economics of poverty, economics of taxation, economics of production, banking and finance …) to evaluate the water pricing policy that is implemented by the water company. These indicators, which are not actually used by the stakeholders involved in water price setting, are relevant for measuring what it is to measure regarding the various performance points of the water pricing policy (connected with the EU-WFD). However, their implementation requires some knowledge of the household water demand functions and also some knowledge of the socio-economic composition of the customer population (facing the water pricing policy that is to assess). This information is precisely provided by an econometric model of household water demand, including the database used for its estimation, provided the latter is tailored to local conditions. In fine, the combination of these two elements forms a Decision Support Tool (DST) that can be used to inform public decision-making on the complex socio-economic consequences of drinking water (and wastewater) pricing, or even to identify optimal policies (for some well-defined decision criteria of local public decision-makers), once coupled with powerful mathematical optimisation software.

The third feature is about the water pricing scheme that the tool is primarily intended to assess, namely the Increasing Bloc Tariff[1] (IBT). IBTs are a pricing scheme that is increasingly used by water managers and public decision makers (in Reunion Island, all the 24 cities make use of this pricing scheme) based on a "social incentive" approach. Basic argument states that IBTs:

- by setting low prices for first cubic meters, enable the households to meet their basic needs at socially acceptable economic conditions (and thus meet the affordability objective set by the EU-WFD);
- by setting high prices for high consumption levels, induce households to adopt water-saving behaviours (and thus meet the "Incentive effect of pricing" objective set by the EU-WFD).

This pricing policy also requires a proper calibration to be financially well-balanced (and meet the "water pays for water" principle), with “taxes” charged on high levels of consumption to fund the subsidies paid on first consumption blocks (and financially balance the cross-subsidy system these pricing policies generate). Academic literature shows then this pricing scheme does not work well in practice, mainly because of poor calibration (which is complex) with sizes of the first subsidised consumption blocks that are often too large.

Some challenges with IBT pricing The problems encountered by water progressive pricing are well known and well identified in the academic literature.

(1) The first refers to the **quality of social targeting**, with the setting of thresholds for the first blocks, which can be considered to meet a social objective (as long as they are subsidized).

Econometric models of domestic water demand provide estimates of required volumes of domestic water to cover basic needs that prove to differ widely among households, depending notably on family size, water-using appliances and consumer habits (García-Valiñas et al. [2010]). A widespread distribution of basic needs for water exists therefore across the population and the point is that, by setting the thresholds of first blocks, an exclusion error is generated with some households who are charged some non-subsidized prices on part of their basic needs. Several studies find that large poor families are over-represented in this category of water users (see for instance Binet et al. [2016]).

(2) The second relates to a **revenue risk for the service funding** (Komives et al. 2005], Foster & Yepes [2006]).

Aiming at limiting the risk of exclusion to which poor large families are exposed, water managers tend to set high values for the thresholds of the subsidized first blocks, with (i) a mass of subsidies that is particularly large and (ii) a contributory part that is based on a small segment of large consumers. With the latter facing particularly high “marginal prices”[2] (and significant increases in their water bills with the introduction of the IBT), the long-term responses of these households may be substantial, with significant reductions in water consumption, what can get them out of their contributory role and, thus, weaken cost recovery (this difficulty also arises when new charges are planned to be funded by large consumers). Besides, the greater the protection afforded to poor large consumers, the greater the risk of revenue for the service funding, since

[1] From a formal point of view, two-part tariffs and linear tariffs can be analysed as specific progressive tariffs, namely a 2-blocs tariff where the consumption threshold above which the unit price of water is increased is (sufficiently) "high". In this way, the MMS is also able to analyse these two pricing schemes.

[2] The marginal price if the price of the block in which the household consumption is located.

small exclusion errors mean a heavy burden for the households who contribute ultimately to the system funding.

(3) **Transfers** Calibration of IBTs requires to target the top of the basic needs distribution to limit exclusion errors. In doing so, this decision makes incurring some large errors of inclusion[3] and the issue of transfers between rich and poor (generated by IBTs) has to be addressed.

To put it simply, if all wealthy families expressed high demands for water and all poor families express low demands for water, the implementation of a proper-calibrated IBT would also result in a transfer of income from wealthy families to poor families (compared to a structurally balanced two-part tariff which is a reference tariff system (see below) that does not generate monetary transfers). This is because wealthy households, by paying a higher price for part of their consumption, given their assumed high demand and the tariff progressivity, would in such a situation contribute to financing the consumption of low-incomes which, because it would be low, would be charged a lower price, subsidised by the collection of a “social tax” charged on high consumption levels (via the cost recovery constraint).

One difficulty, however, is that when it comes to water consumption, poor households are not as different from wealthy households as one might initially think. Nauges and Whittingthon [2017], by combining data from several sources, find a positive but quite low correlation between household water use and income, what means "that there are many rich households that use small amounts of water, and many poor households that use large quantities of water" (the quality of water-consuming equipment owned by households clearly plays a role here). In this respect, since IBTs implement transfers from large to small water-user households, they also entail transfers from low incomes to high incomes. Several studies demonstrate then that IBTs perform poorly regarding the subsidy targeting for low-incomes with high-incomes that receive a large share of subsidies[4]. In fine, progressive tariffs may thus increase income inequalities, compared to a structurally balanced two-part tariff, with a system of cross-subsidies that would be anti-redistributive.

(4) **Progressive pricing is not as much of an incentive as it should/could be**. As emphasized by Monteiro & Roseta-Palma [2011], there is no particular reason, from a theoretical point of view, to manage the incentive effect of water pricing through progressivity. However, it is understood that it is always possible, by setting sufficiently high prices for excessive levels of consumption, to induce large consumers to use the resource more sparingly (Wichman [2014]). Further, some empirical studies show that price sensitivity of water demand is higher when consumers face IBT (Cavanagh et al. [2002]), so that the choice of this pricing scheme may indeed generate more attention from the households on their water uses.

However, empirical evidence suggests as well that progressive pricing is not as incentivizing as it should be (Gaudin [2005], Martins & Fortunato [2005], Pérez-Urdiales et al.[2022]), with consumers who tend to misperceive IBTs. Faced with a tariff that presents a certain complexity, they tend in particular to think in terms of average price (Liebman & Zeckhauser [2004], Cartter & Milon [2005]), what can lead them to over-consume with an under-estimation of the "marginal price" (that is, the unit price of the block in which the household's consumption is located) that

[3] Inclusion error refers to social pricing on part of consumptions that do not meet basic needs.
[4] See notably Gómez-Lobo & Contreras [2003], Komives et al. [2007], Barde & Lehmann P. [2014] and Fuente et al. [2016] and Whittington & Nauges [2020].

has to be considered for a fully optimal management of domestic uses (as basic consumer theory may indicate). In this way, an IBT may thus lead to a reduction in overall consumption when the tariff is properly perceived, but to an increase in overall consumption, given the overconsumptions that are generated, when the tariff is poorly perceived.

It should also be borne in mind that the reduction in overall consumption, when switching from two-part tariff to social progressive pricing, is not a granted output. Implementation of a social (and presumed) incentive IBT may indeed lead to an increase in overall consumption, even though the unit price schedule were fully well-perceived (Crampes & Lozachmeur [2014], Mayol A. & Porcher S. [2019]). Given the context of the decision problem, this type of IBT subsidizes indeed the small consumers and taxes the large consumers with a net effect on overall consumption that depends on the calibration of the tariff (Ito [2014], Mayol A. [2017]), with the sizing of the first subsidized consumption blocks, the level of the various subsidy rates (per cubic metre) for each of these first consumption blocks, and the level of the (endogenous) unit margin rates for the higher consumption blocks to complete the financing of the cross-subsidy system.

The point is that academic literature is able to make these observations and also identify some ways of improvement precisely because it applies econometric demand models (commonly used by economists) and appropriate indicators for the various items to be analysed. In addition, a number of factors suggest that water managers and local public decision-makers, when considering the content of their pricing policies, may lack a clear understanding of the socio-economic consequences of the decisions they are making, or may even be basing their thinking on implicit assumptions about key factors that are not actually met (such as the fact that small consumers are mainly low-income earners and large consumers are high-income earners). This lack of clarity can then lead to decisions which, although well-intentioned, may turn out to be ineffective or even counterproductive in terms of the objectives pursued.

The development and dissemination of the microsimulation model is intended to respond to this lack of clarity for stakeholders, by enabling them to establish a clear diagnosis of the gaps on each of the performance points related to the EU-WFD and, above all, to inform them of the nature and extent of the trade-offs between the various objectives the water pricing policy has to meet in fine. As apparent, it is here to integrate economic and social analyses into decision-making processes, and "break boundaries between services valuation [...] and the employment of economic and social sciences" (Water JPI 2018 Joint Call).

Outline The content of this document is set as follows.

Section 1 presents the general architecture of the tool, which is divided into 5 modules: the Population module, the Tariff module, the Demand module, the Invoice module and the Evaluation module (dashboard).

Section 2 briefly describes the data in the Population module (taken from a survey conducted in 2005 on a representative sample of households living on Reunion Island) on which the tool operates (for demonstration purposes).

Section 3 presents the tariff module in which the user enters the values of the IBT tariff parameters to evaluate, as well as some primitives of the decision problem he faces such as the production costs of the drinking water supply, the production costs of domestic wastewater collection service, the environmental costs, the tax system with VAT and charges for the protection of the aquatic environment, and some social data such as the poverty line.

It should be emphasized that the construction of this module is based on the French regulations governing the pricing of drinking water and wastewater services. These include the principle of equality of public service, what means that the pricing system must apply to all customers, regardless of the level of their income (universality of service). In addition, the public collective wastewater service is also billed by the company in charge of managing the service, according to a specific tariff, if the household is actually connected to the network (which concerns slightly more than half of the households in the Population module) and is not billed otherwise. Section 3 also provides some elements of analysis related to the pricing of drinking water supply such as the notion of structural progressivity, Nordin's D and "Two-part Structurally Balanced Tariff" (abbreviated as **TBSE** for "Tarification Binôme Stucturellement Equilibrée" in French) which is a natural point of reference for the pricing of public drinking water and wastewater services.

Section 4 is devoted to presenting the Demand module and the econometric model estimated by BCP [2014], which is used by the tool to calculate and break down household water consumption (for each household subscriber in the Population module). As the user has the possibility of modifying the values of the response coefficients (that correspond, by default, to the ones estimated by BCP [2014]), it is important for the user to be aware of what these numeric values measure (the latter have an economic significance) and, more generally, of what a demand function represents. To this end, section 4 provides information on the main empirical and theoretical properties of household water demand functions, the particularities of the BCP specification (with which the tool works), and the algorithms used to compute and decompose household water consumptions.

Section 5 presents the Invoice module which is the equivalent of a water account for each household (subscriber) in the Population Module, including the levels of its water consumption, captive consumption, basic consumption, non-basic consumption, overconsumptions (related to tariff misperception) …, and the identification of the various impacts exerted by the pricing policy (which is evaluated/tested by the user) such as, for instance, the amounts of the gross and net subsidies that are granted to a household by the IBT (around sixty variables are identified in this file).

Sections 6 to 10 give a detailed presentation, with one section for each field of analysis listed below, of the indicators used by the tool to measure the multi-dimensional performance of the water (and wastewater) pricing policy which is evaluated/tested by the user. **Section 6** deals with affordability, based mainly on two basic indicators: the CAR (Conventional Affordability Ratio) that relates to the proportion of income that a household spends on paying its water bill and the PAR (Potential Affordability Ratio) that relates to the proportion of household income spent on water bills to cover basic needs. **Section 7** deals with the incentive effect of pricing, in particular by measuring the gains and losses of the IBT in terms of household water consumption (compared with the TBSE), the measurement of over-consumption (linked to IBT misperception) and the associated private costs of poor management (borne by household), and the measurement of inclusion and exclusion errors, in volume terms, linked to tariff calibration (and the setting of consumption thresholds). **Section 8** deals with the issue of equity (measurement of cross-subsidies, quality of social targeting, impact of the IBT on household income inequalities …), **section 9** with economic efficiency (surplus analysis) and **section 10** with cost recovery (and characterisation of the financing structure).

In terms of the information provided, it should be noted that the tool always starts by giving some general information, for the population as a whole, with basic descriptive statistics that are:

- Measures of "Central tendency": Average, Median,

- Summary of the Distribution Function ("box plot ") : Min (minimum observed value), Max (maximum observed value), Q1 (first quartile), Q3 (third quartile), D1 (first decile), D9 (last decile), F_Mean (the percentage of statistical units (usually households) below the mean),

- Dispersion (Heterogeneity) : Variance, Standard deviation, MAPE (mean of deviations from the mean), Coefficient of variation, Interquartile range, Interdecimal range,

- Symmetry / Asymmetry: Yule coefficient,

- Concentration: Gini index, Schutz index, inter-deciles ratio, inter-decimes ratio, "S80 / S20" ratio,

for the various variables of interest. On this basis, it focuses on a breakdown of the population of households (subscribers) into sub-populations, usually by breaking down households into 2 groups according to:

(i) whether the household customer is connected to the collective sewerage network (in which case he faces and pays the Drinking Water and Wastewater tariff) or not (in which case he faces and pays the Drinking Water tariff only),

(ii) whether their standard of living is below the poverty threshold entered by the user (in which case they are part of the Poor Household group) or not (in which case they are not part of the Poor Household Group)

or even into 4 groups (with the intersection of these two criteria) and 10 groups (with the household population broken down by consumption deciles or by standard of living deciles). presenting the data by consumption deciles or by household standard of living deciles. The MMS can also make occasional use of specific tools (ROC space, Pen parades, relative benefit distribution curves, absolute beneficiary curve, Gini index decomposition ...) and specific indicators (Sen index, Schutz index, Omega ratios, leakage rates, contributions to aggregate (or social) surplus, consumer surplus ...). As the latter are not necessarily familiar to non-specialists, this document presents in some detail the construction of several of these elements (see also the catalog of variables provided at the end of the docuement).

Section 11 presents the "scaling up" operation of the basic model, the calculation and spatial breakdown of some indicators that are used in the basic model, and some of the new indicators that could be used to measure spatial inequalities that one may wish to measure when it is to assess a set of water and wastewater pricing policies carried out in a dedicated geographical area, such as the catchment area.

Section 12 concludes with a discussion of the added value of the tool, some points on which further development could be considered, and the necessary conditions to exploitation and replication. On this matter, it should already be emphasised that the tool to be truly operational (and provide effective information on the performance of a specific pricing policy implemented in a well-defined geographic area) needs to substitute the econometric model estimated by BCP [2014] with an econometric model of local household water demand (including the data file used to estimate this econometric model of local demand) and, if necessary, to adapt the Tariff module to the specific features of national regulations. As it stands, the tool can nevertheless be used for demonstration and training purposes, whether for students of economics, management of public organisations or water sciences, or for stakeholders willing to enhance their skills in the socio-economic dimension of household water demand management policies, and pricing policy in particular. It can also be used by members of the general public, particularly those wishing to find out more about the public debate on water pricing, or to support initiatives aimed at encouraging Citizen Engagement (notably by providing some general or personalized information on the general and personal impacts of the water pricing policy that is or could be implemented).

The code The tool is a web application developed in Python with Flask for the back-end and Angular for the front-end. Flask is a python framework allowing the rapid integration of the logic of a python writing simulator. Angular is a front end framework that allows the user to have the neat interactivity required in the simulator settings and results visualization workflow.

These choices mean that the tool can be used in two different user cases. In the first one, the user uses the version that is hosted by a third party server (in this case the tool can be considered as a SAAS (software as a service)). Therefore, there is no need to install the source code nor the software needed to run it (Python, Javascript, Angular), and the tool can be used like any web application on a web browser. In the second case, the user has more advanced knowledge and can host the tool on its own computer, to allow for more transparency and flexibility. In this case, the user has the possibility to consult how the calculations are made, or to modify the algorithms to match its own needs. This configuration also allows to easily decouple the front-end and the back-end or even to use the latter as a tool to perform stand alone processing. Indeed, the back-end is a collection of algorithms that can be used directly as a Python library or as microservices connected to a front-end.

The choice of Python over R has been made to allow as many users as possible to tailor the tool to their needs by modifying existing modules or developing additional modules (although algorithms can be transcribed back into R and it is also possible to use both languages by creating an API).

The execution flow of the simulator has been implemented using two design patterns that will facilitate its evolution and adaptability. Each matrix is processed through a sequence of steps, implemented using a design pattern known as the "chain of responsibility." This execution flow can thus be customized by inserting one or more additional steps. Furthermore, a second design pattern, known as the "factory" pattern, is used to build execution chains based on specific needs. The combination of these two design patterns provides a valuable level of flexibility for adapting the simulator to different constraints and contexts.

The use of the tool will be governed by a gpl-3.0 licence (filed during 2025) which allows the code to be opened up under the following two obligations: (1) mention of the InnWater project and (2) sharing, by making them public, of any additions and modifications. The author of this report would like to extend his warmest thanks to Pedyl Maree and Valentin Morin, both of whom worked on the development of the code during research placements carried out as part of the Newts project (Water JPI 2018).

II – GENERAL ARCHITECTURE

The MMS includes 5 modules:

Population module describes the population of households with regard to some relevant variables for calculation and decomposition of (i) water consumptions, (ii) water bills and (iii) indicators making up the dashboard. Unless otherwise stated, households are indicated by the letter *i* (with data that has been anonymised).

Pricing module describes the tariff system (based on French regulation) that applies for the households listed in the Population module. In the case of IBTs, this includes: (i) fixed part; (ii) number and size of consumption blocks; (iii) unit prices within each block; (iv) VAT rate; (v) environmental charges for (domestic) water utility and (domestic) wastewater utility. The user is also asked to provide information on service costs (fixed charges, number of domestic subscribers, unit variable cost of production for, successively, domestic water supply and domestic wastewater supply), environmental costs (linked to domestic uses of water) and social data (like the poverty line).

Demand module makes use of an econometric model to compute for each household in the Population module:

- its water consumption $q_i = q_i^d(\cdot)$ where q_i is the water consumption of household *i* and $q_i^d(\cdot)$ its water demand function,
- the related captive part noted q_{0i},
- the size of its basic consumption noted $\underline{q}_i$, like for instance:

  $$\ln \underline{q}_i = -2.56 + 0.48 \ln N_i + 0.44 \cdot \mathrm{SNWA}_i \quad (2.1)$$

  (captive part is reprocessed by the user to estimate basic water needs; see section 4) where N_i denotes the size of family *i* and SNWA is the **S**hare of **N**on-Working **A**dults, computed as the ratio of the number of inactive adults to the number of adults (non-employment rate), within family *i*,

in interaction with the Population module (that provides the necessary socio-economic information) and the Pricing module (in which the IBT the user is intended to assess is set).

Invoices module collects all relevant information at the household level with the setting of a specific bank account for "Water and Wastewater expenditures". This includes:

- the computation of the water bills $T_i = T(q_i^d(\cdot))$ spent on water and wastewater utilities by household *i* (given the tariff parameters that have been set by the user),
- the breakdown of the water bills T_i firstly into a captive part $T_{0i} = T(q_{0i})$ and a non-captive part $T_i - T_{0i}$, next into a minimum basic part $\underline{T}_i = T(\underline{q}_i)$ and a non-basic part $T_i - \underline{T}_i$ ... for water (and wastewater when the household is connected to the sanitation network),
- the accounting reconstruction of invoices with the identification of various subsidies and contributions to service funding with reference to the "structurally balanced two-part tariffs",

- the measurement of public and private welfare gains and losses, in monetary equivalents ("aggregated surplus ", "consumer surplus"), generated by the implementation of the IBT that is set/tested by the user
- ...

for each household in the Population module.

Evaluation module includes a scoreboard and about sixty indicators (for level 1 information) allowing to measure the socio-economic performance of the pricing policy in each of the 5 fields of analysis related to EU-WFD with: Affordability; Incentive Effect; Economic Efficiency ("Welfare"); Equity; Cost Recovery (Quality of the Funding)). These indicators correspond (i) to those commonly used in academic literature on water; (ii) to other measures commonly used in other fields of Economics (Poverty Economics, Production Economics, Tax Economics, Banking and Finance ...) that are thought to be relevant for the field of analysis considered.

As regards the use of the tool, its handling and the interactions between the various modules, the following points should be noted.

Primitives Initially, the user is asked to declare (in an extract from the Tariff module) (1) the production and distribution costs of the drinking water service with :

- the amount of fixed costs noted CF_{EP} (the acronym EP stands for Eau Potable in French);
- the unit variable cost $\text{CVM}_{\text{EP}} = c_{\text{EP}}$ (the acronym CVM stands for "Coût Variable Moyen" in French), which is assumed to be constant (this variable therefore gives the cost of producing and distributing one cubic metre of drinking water, excluding fixed costs);

as well as (2) the production and distribution costs of the (collective) wastewater treatment service with :

- the amount of fixed costs noted CF_{A} ;
- the unit variable cost CVM_{A} , which is (also) assumed to be constant (this last variable therefore gives the production and distribution cost of one cubic metre of wastewater / after use by the household, excluding fixed costs).

Users are also asked to enter :

- the number of domestic subscribers (households) to the public drinking water service, noted as n (sometimes n_{EP});
- the number of domestic subscribers (households) to the public sewerage service, noted n_{A}, with $n_{\text{A}} \leq n = n_{\text{EP}}$;

and to answer a question relating to the processing of the data ("sample adjustment"), which is described in detail in the next paragraph. On this basis :

- The number of domestic subscribers not connected to the public sewerage network (in number $n - n_{\text{A}}$) forms a specific group of Subscribers that is denoted G1. The size of this group, noted n_1 with $n_1 = n - n_A$, then determines the number of households benefiting only from the public drinking water service.

- The number of domestic subscribers connected to the public sewerage network, n_{A}, forms a specific group of Subscribers known as G2. The size of this group, denoted n_2 with $n_2 = n_{\mathrm{A}}$, determines the number of households benefiting from the public drinking water and wastewater service, known as the "EPA" service (this acronym stands for "Eau Potable et Assainissement" in French).
- The provision of the drinking water service to the $n = n_{\mathrm{EP}}$ household in the Population module and of the collective wastewater service to the $n_2 = n_{\mathrm{A}}$ households in Group 2 then form the general service known as the "EP / EPA" service (this acronym stands for "Eau Potable / Eau Potable et Assainissement" in French).

These values (production and distribution costs for "EP" service and "A" service, number of household subscribers to "EP" service, number of household subscribers to "A" service) entered by the user set the parameter values for the structurally balanced Two-Part Tariffs "TBSE" (this acronym stands for "**T**arification **B**inôme **S**tructurellement **E**quilibrée" in French), the TBSE "EP" and the TBSE "A", which are the reference tariffs for the public drinking water service and the public wastewater service respectively. Excluding charges (collected by the local water agency) and VAT (collected by the Operator for the benefit of the State), these pricing schemes are made up of:

- a subscription fee (also known as the fixed part), noted F ;
- a constant price per cubic metre, noted π ,

(the letter *p* is reserved for designating another variable) with (i) a subscription amount equal to the amount of fixed costs divided by the number of subscribers :

$$F_{\mathrm{EP}} = \frac{CF_{\mathrm{EP}}}{n}$$

$$F_{\mathrm{A}} = \frac{CF_{\mathrm{A}}}{n}$$

and (ii) a price per cubic metre equal to the value of the average variable cost (entered by the user):

$$\pi_{\mathrm{EP}} = c_{\mathrm{EP}}$$

$$\pi_{\mathrm{A}} = c_{\mathrm{A}}$$

based on a quarter of consumption (and a quarterly billing period). The user also enters :

- the VAT rates (t_{EP} and t_{A}) for the drinking water and wastewater services;
- the values of the charges, noted r_{EP} and r_{A}, in euros per cubic metre (excise duty) for the drinking water and wastewater services;
- the value of the environmental cost, noted c_e, in euros per cubic metre, defined as the cost of completely depolluting one cubic metre of (domestic) waste water;

as well as certain 'social' data with :

- the value of the poverty line (for the (modified) OECD equivalence scale)

- the threshold value for the CAR (weight of the "EP / EPA" bill in household income) above which the user considers the household is facing an affordability issue in the CAR sense;
- the threshold value for the PAR (weight of the EP / EPA bill in the household's income to cover its basic needs) beyond which the user considers the household is facing an affordability issue in the PAR sense.

Following these initial entries, the user is also prompted (in an extract from the Demand module):

(i) validate or modify the default values entered for the parameters (response coefficients) of the demand function ;

(ii) validate or modify the list of determinants (of the demand function) used to calculate basic consumption (of households in the Population module).

All this data is then saved in the "My Simulations" portfolio of simulations for automatic loading of these parameters.

Initialisation On this basis, and after the user has run the algorithm, the tool displays (1) a simplified aggregate dashboard, made up of around twenty indicators, providing information on :

- the affordability of TBSE (in the general population and in poor households),
- TBSE consumption levels (which can spontaneously be relatively high) for drinking water and wastewater services,
- deviations in consumption compared with the so-called first-order social optimum (with full cost recovery, including the environmental cost) and the accompanying losses in aggregate surplus,
- some financial data (including the breakdown of sales for the general "EP/EPA" service),

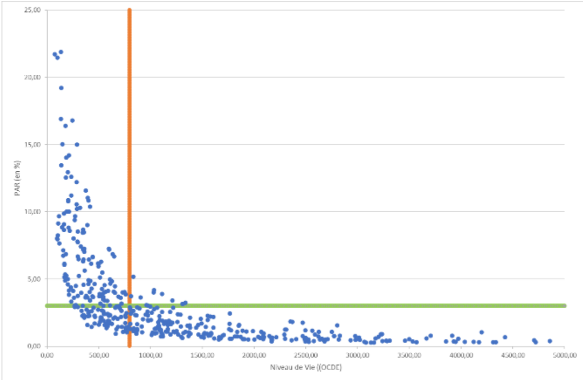
and 4 computer graphics (see Figure 1, page 25):

(1) TBSE Affordability: a scatter plot with (i) on the x-axis, the standard of living of households; (ii) on the y-axis, the value of the TBSE PAR (as a reminder, the proportion of income that a household devotes to meet its basic water needs) and (iii) a zoning system enabling to identify households (including poor households) facing an affordability issue or in a situation of vulnerability (i.e. close to water insecurity, as defined by the PAR);

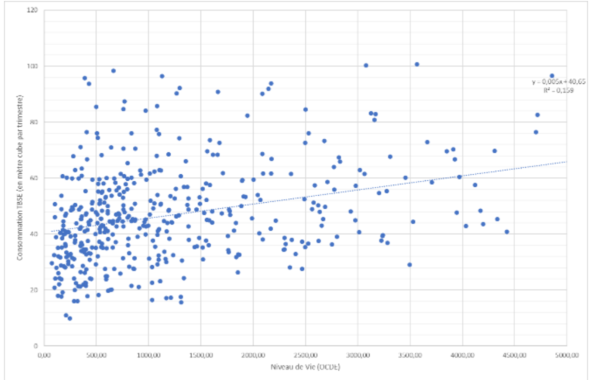
(2) Distribution of Basic Needs: the Pen's parade of basic consumption (which provides useful information for setting the sizes of the first consumption blocks that are subsidised with progressive pricing of the "social incentive" type) and the Pen's parade of captive consumption (which determines a minimum level of service and, in so doing, a minimum value for the cost of the service);

(3) Correlation between consumption and standard of living: a second scatter plot with (i) household standard of living on the x-axis, (ii) TBSE consumption on the y-axis and (iii) some statistics on the degree of correlation (covariance, linear correlation coefficient, Spearman coefficient (rank correlation)) between these 2 variables;

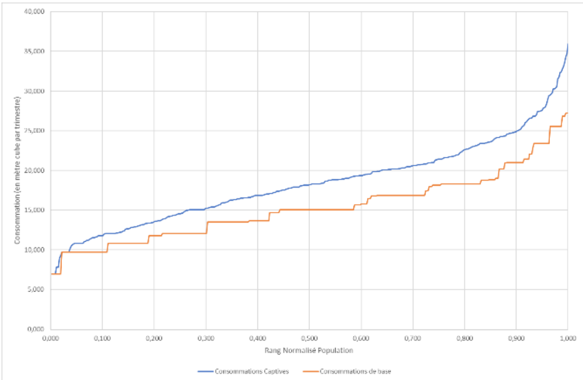
Figure 1 : Initialisation - TBSE



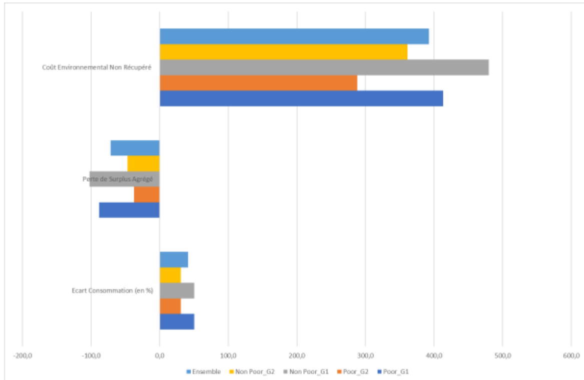
1.1: TBSE PAR Affordability and standard of living



1.3: TBSE Consumption and living standards



1.2: Pen's Parade of basic and captive consumptions



1.4: Consumption deviation, Welfare losses and environmental cost recovery

(4) TBSE deviation from the first-rank optimum: a bar chart showing the average values of the deviations, in percentage points, of consumption from their first-rank values, the related losses in aggregate surplus and the amount of the environmental cost that is not recovered in the general population and for different categories of Household (G1-Poor, G1-Non Poor, G2-Poor and G2-Non Poor).

It is also displayed (i) as a reminder, the value of the price elasticity (which has been validated/entered by the user) and (ii) some basic statistics on the distribution of income elasticities.

First round evaluation Based on these initial elements, the user is invited to declare the parameters of the new tariff in the Tariff module (Two-part Tariff or Increasing Block Tariff that he intends to test and evaluate) for the drinking water tariff and the wastewater tariff successively, after which the tool displays :

- the amount of the (likely) subsidy on the Right of Access for the "EP" service, the "A" service and the "EPA" service;
- the subsidy rates, in euros per cubic metre, for the various subsidised consumption blocks;
- the "tax" rates, in euros per cubic metre, for the different consumption blocks that are "taxed" (marked);
- an infographic identifying, for the "EP" and "EPA" services, (i) the consumption level at which the margin generated on the Household's consumption becomes positive and (ii) the Household break-even point, defined as the consumption level at which the Household contributes to the financing of the service Access Fee Included or "DAI" (this acronym for "**D**roit d'**A**ccès **I**nclus" in French).

Once the user has validated the pricing parameters, the tool proceeds to:

- the calculation of domestic water consumption (for each household in the Population module),
- the breakdown of this domestic consumption into the sum of 4 elements: basic consumption, captive but non basic consumption, a variable part (excluding over-consumption) and over-consumption (possibly equal to 0) linked to a poor perception of the tariff,

and a number of complementary treatments, including :

- the breakdown of the bill into the sum of 4 elements, linked to the breakdown of domestic consumption (including the expenditure borne by the household to cover its basic consumption and an additional expenditure linked to its over-consumption),
- the calculation of the gross subsidies granted and the gross "taxes" (margins) levied by the operator on the Right of Access and per consumption block,
- the calculation of net subsidies and "net taxation" (in fact, contributions to service funding) Acces Fee Excluded or "DAE" (this acronym for "**D**roit d'**A**ccès **E**xclu" in French), and Access Fee Included or "DAI",
- the calculation of (gross and net) subsidies and (gross and net) "taxes" on basic consumption and on the supply of basic service (which includes the Right of Access) generated by the "EP" tariff and the "A" tariff ...
- the calculation of VAT amounts ...

- the calculation of the percentage of households facing an affordability issue, the calculation of the affordability deficit …
- the calculation of the impact of the IBT on the level of service (which may be higher than the TBSE), the calculation of the mismanagement costs that are borne privately by households ...
- the calculation of the inclusion and exclusion errors in volume and in value ...
- the rates of coverage of the fixed costs of the service (in the economic sense of the term) by the variable part (in the economic sense of the term) of the sales revenue ...

to feed a simplified dashboard, known as the "First-Round Dashboard", made up of around sixty indicators enabling the IBT's performance (and its added value compared with the TBSE) to be measured in the first instance. The latter is accompanied by the four previous infographics for viewing some key-points relating to the performance of the IBT under consideration (except for a slight modification to the second diagram, which no longer shows the Pen parade of captive consumption but the distribution function for the mass of basic consumption, in addition to the Pen parade of basic consumption[5]). On these particular points, see (1) the table " able 1 : Agregate Dashboard – IBT vs. TBSE (first round)", page 28, and (2) the set of figures "Figure 2 : First Round IBT (vs. TBSE)", page 29, for illustrative purposes.

Depending on the user's assessment of these results, the user can then go back to calibrating the pricing policy (to adjust it) or, if the initial results obtained seem worthwhile, validate the EP and A tariffs entered at the start of this first phase, and launch a full assessment.

Complete Evaluation The tool provides a comprehensive analysis of the socio-economic performance of the user's tariff policies in each of 5 main areas of Affordability, Incentive Effectf, Equity, Economic Efficiency and Cost Recovery (quality of funding). To this end, (1) some additional indicators are calculated and (ii) some focus on particular groups (such as poor households, for example) are performed, and, for the appropriate indicators, (3) the links between the figures obtained for the General Population and those obtained for the various sub-populations (of the General Population) are highlighted, with the relevant breakdowns. Last but not least, (4) the performance of pricing policies on some points of vigilance, in relation to the difficulties identified in the academic literature, are checked. In order to limit calculation time and, above all, to respect the logic of the various questions on which this evaluation process is based, these different operations are carried out in stages, for each of the major fields of analysis, with an increasing degree of granularity of information. Finally, users can export their data for further processings that are not implemented with the current version of the tool.

[5] By way of illustration, the diagram informs that 40% of households have a basic consumption of less than 16.8 cubic metres per quarter (this information is provided by the Pen's parade of basic consumption) and that the basic consumption of these 40% of households (whose basic consumptions are the lowest ones) represents 29.7% of the total basic consumption (this information is provided by the distribution function of the mass of the basic consumptions). Alternatively, the diagram also shows that subsidising the first 20 cubic metre of water consumption, through the setting of the threshold of the first subsidized consumption blocks at 20 cubic metre per quarter, enables to fully support the basic needs of 86.5% of households, with the corresponding 79.6% of the total basic consumption that is fully subsidised (it should be noted that this rate is, by construction, lower than the percentage of total basic consumption that is subsidised, as the $100-86.5=13.5\%$ of the subscriber population who are not subsidised for all their basic consumption are subsidised here for the first 20 cubic metres).

able 1 : Agregate Dashboard – IBT vs. TBSE (first round)

Affordability (PAR H)		
	IBT	TBSE
Headcount ratio	15.9%	32.1%
App. Afford. Deft	3.37 €	16.87 €
Effec. Afford. Defit	17.69 €	53.33 €
Gini_App	0.956	0.793
Gini_Eff	0.725	0.355

Incentive Effect		
Average Consumption (m3 / trim)		Average Bill €/trim
IBT	40.9	91.6
IBT_PP	38.5	73.8
TBSE	47.5	95.8
Eff Overconsumption		Eff Mis-mng Cost
per H	6.7	19.8
per Ind	2.3	2.3

Economic Efficiency		
Average	Conso	Delta W
	m3 / trim	€ / trim
First Best	90.4	***
Delta IBT PP	-51.9	-961.28 €
Impact Sur_Co	2.4	198.03 €
Delta TBSE	-42.9	-170.46 €
Delta Surplus M		-781.05 €

Equity	Net Income Gini Index	
IBT	0.491	
IBT-AE	0.491	
TBSE	0.494	
	DAE	DAI
Net Sub Basic C	0.09 €	41.65 €
Omega ratio	0.67	0.97
Net Taxes Basic C	9.43 €	0.00 €
Omega ratio	1.26	***
	AFE	AFI
Net Sub C	0.00 €	12.18 €
Omega ratio	2.12	1.31
Net Taxation	50.98 €	12.17 €
Oméga ratio	0.85	0.58

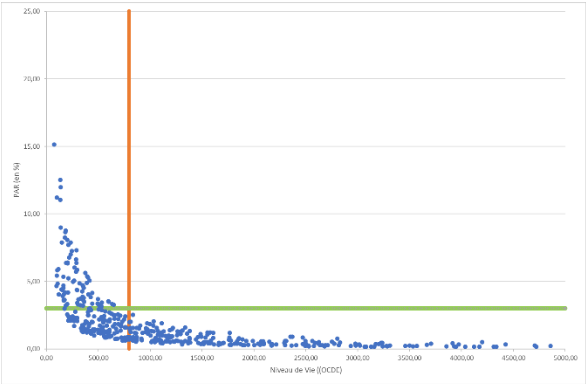
Funding	Général	% Total Cost
REX_Op	**-179 449 €**	-0.8
	% H	
Net Contributors	42.8	
Net Beneficiaries	57.2	
Subsididized basic C (en %)		
	69.20	
Subsididized non basic C (en %)		
	41.6	
Margined C (en %)		
	34.5	
	In %	
"Bad" Sub	0.1	
"Bad" Tax	18.2	

Unrecovered Environmental Cost		
		m3 / trim
TBSE Conso Rang 1		423.18 €
TBSE		683.76 €
IBT		192.41 €
IBT_PP		181.08 €

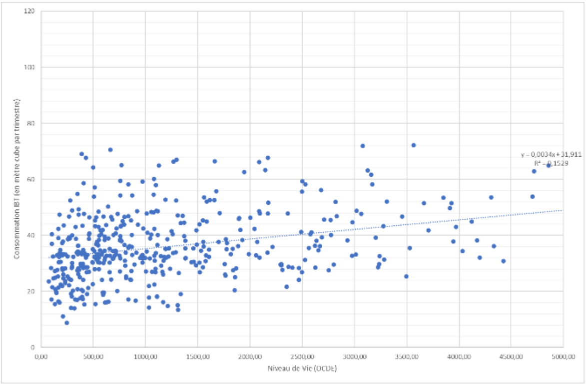
Water Agency		Total annuel
Excise duty		957 154 €

State		Total annuel
VAT		1 076 968 €

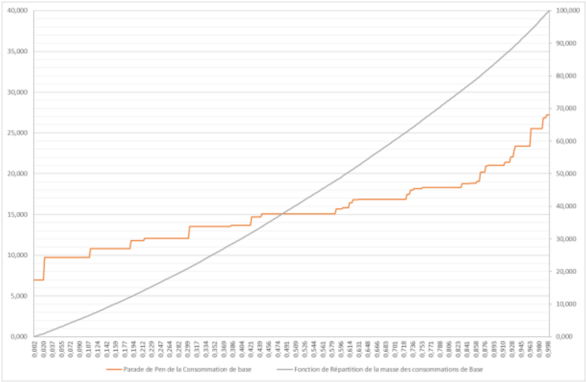
Figure 2 : First Round IBT (vs. TBSE)



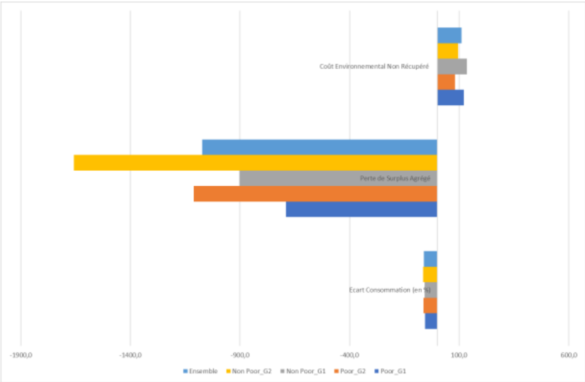
2.1: IBT PAR affordability and standard of living



2.3: IBT consumption and standard of living



2.2: Captive consumption - Pen parade and distribution function



2.4: Pression, Perte de Bien-Être et Récupération Coût Environnemental

III – POPULATION MODULE

The Population module contains a Household data file describing the socio-economic characteristics of a sample of the population living in Reunion Island. These data are extracted from a database built by Binet et al [2014] as part of a research contract with DIREN (DIrection Régionale de l'ENvironnement, now DEAL-Réunion) aiming at identifying the determinants of domestic water consumption in Reunion Island. It comprises 15 variables, with $n = 458$ entries per variable, listed below:

Climatic data: frequency of number of rainless days during billing period (quarterly).

Household characteristics: household income, imputed household income (following an imputation method detailed in Carlevaro et al. [2007]), household size (number of people), number of children, number of working adults, number of non-working adults, number of equivalent adults (modified OECD scale).

Equipment characteristics: Jardin (dummy variable which takes the value 1 if the household is equipped with a Garden, 0 otherwise), Piscine (dummy variable which takes the value 1 if the household is equipped with a Pool, 0 otherwise), Assaini (dummy variable which takes the value 1 if the household is connected to the collective sewerage network, 0 otherwise), NonAssaini (dummy variable which takes the value 1 if the household is not connected to the collective sewerage network, 0 otherwise), Maison (dummy variable which takes the value 1 if the household lives in a detached house, 0 otherwise), Retraité (dummy variable which takes the value 1 if the household is retired, 0 otherwise), Proprio (dummy variable which takes the value 1 if the household owns its home, 0 otherwise).

It should also be noted that, compared to the original data set, the income series has been increased exogenously by 15%.

The tool then assumes, for demonstration purposes, that this sample is representative of the population of households residing in the area the water manager is in charge or, more precisely, (i) that the sub-sample of households connected to the sewerage network (so-called Group 2 households) and (ii) that the sub-sample of households not connected to the sewerage network (so-called Group 1 households) are representative of these two segments of the customer population in the area the manager is in charge of. Based on this assumption, the tool constructs a representative sample of the $n = n_1 + n_2$ subscriber population of 1,000 (default value) or 2,000 or 3,000 … households by drawing lots (i) from the BCP-Group 1 sample, a number of households equal to $1000 n_1 / n$ for the "User Group 1 sample", and (ii) from the BCP-Group 2 sample, a number of households equal to $1000 n_2 / n$ for the "User Group 2 sample", based on the information entered by the user for the number of subscribers $n = n_{\mathrm{EP}}$ and $n_2 = n_{\mathrm{A}}$ with the description of the cost functions for the EP and A services. Next, the tool creates variability in the data by noising, for each draw, the income of the household that is drawn with the addition of a realization drawn in a lognormal distribution of mean zero and variance 0.01 (the user is given the option of modifying the value of this variance). In parallel with this procedure, the user also has the option of loading into the Population module a data file that is assumed to be representative of the subscriber population (and consistent with the information entered on production costs) for the area under consideration.

IV – TARIFF MODULE

Preamble In the following, it is noted:

- q the household water consumption, q_i the water consumption of Household i,
- T the amount of the bill for the drinking water supply (referred as the EP service) or for the drinking water and wastewater supply (referred as the EPA service[6]) according to the household is connected to the collective sanitation network or not,
- $T = T(q)$ the tariff function that links the amount of the bill to the household consumption (volumetric pricing).

Where appropriate, one can also write $T = T(q, \theta)$ with θ a vector of tariff parameters (characteristic of the tariff that is evaluated/tested by the user). What follows is a presentation of the two main tariff schemes, Two-Part Tariff and Increasing Block Tariff (IBT), which the tool aims to assess. In order to simplify the presentation, taxation will be initially ignored.

4.1 Two-Part Tariffs

By definition, a two-part tariff of parameters $\theta = (F, \pi)$ is a pricing policy for which:

$$T(q) = F + \pi q \tag{4.1}$$

where:

- F is the amount of the subscription fee, which is interpreted as an access fee to the EP / EPA service (fixed cost of consumption) that the subscriber (household) has to pay to start consuming tap water;
- π is the unit price, per m^3.

This pricing scheme includes the following special cases:

- Flat-fee pricing:

$$T(q) = F \tag{4.2}$$

(in this case, $\pi = 0$ and pricing is no longer volumetric);

- Linear pricing:

$$T(q) = \pi q \tag{4.3}$$

(in this case, $F = 0$).

- Structurally balanced two-part tariff:

$$T(q) = \frac{CF}{n} + c \times q \tag{4.4}$$

denoted also TBSE (for "Tarification Binôme Structurellement Equilibrée" in French) with n the number of subscribers (households), CF the level of fixed costs (borne by the operator) and c the unit variable cost / the marginal cost of production (assumed to be constant).

[6] The acronyms EP and EPA stand for "Eau Potable" and "Eau Potable et Assainissement" in French.

As a reminder:

- with some exceptions, flat-fee pricing is prohibited in France;
- linear pricing corresponds to the "first best" pricing when $\pi = c$ (marginal cost pricing) with an operating deficit (borne by the operator) equal to the amount of fixed costs CF (in this scenario, the service is partly financed by taxation with a transfer from public authorities to the operator, most often as part of a planning contract);
- the TBSE $(F, \pi) = \left(\frac{CF}{n}, c\right)$ corresponds to a "Ramsey-Boiteux" type pricing, when the operator is subject to a cost recovery constraint and water pricing has to satisfy the "water pays for water" principle.

In this respect, TBSE constitutes a reference tariff from which the effects linked to:

- a change in tariff parameters F and π such as, for instance, the capping of the fixed part introduced by some national regulations[7], including the abolition of the fixed part ($F = 0$) for which several user associations are campaigning[8],

or:

- a tariff reform such as the one announced in March 2023 by President of the French Republic, that will extend progressive pricing to the whole of France [9]

can be measured. It should also be noted that any two-part tariff for which $F < \frac{CF}{n}$ subsidises the access fee with a unit price π that has then to be greater than the unit variable cost $\mathrm{CVM} = c$ to ensure the financial equilibrium of the service / when the "water pays for water" principle applies.

Example By way of illustration, Figure 3 shows the pricing graph for the EP service of the city of Melun-Sénart 77 with $F = 9.2629$, $\pi = 1.8389$ and $T(q) = 9.2629 + 1.8389 \cdot q$ for one consumption quarter (billing period). The horizontal line (drawn in red) gives the unit variable cost of consumption which is borne by the household. The latter is also equal to the cost of one additional unit, referred as the “household marginal cost of consumption”, with:

$$T'(q) = \pi = 1.8389$$

As shown in Figure 3, one feature of the two-part tariff system is that it specifies a household marginal consumption cost that is constant/does not depend on the household's level of consumption (flat rate).

7 In France, subscription fee can not exceed 30% of the bill, calculated for a typical consumption of 120 m³ (40% in rural municipalities). This restriction does not apply to municipalities classified as "tourist".

8 See for instance "Water prices: the pricing structure can penalise small consumers", https://www.clcv.org/communiques-de-presse/prix-de-leau-la-structure-tarifaire-peut-penaliser-les-petits-consommateurs.

9 https://www.francetvinfo.fr/meteo/secheresse/direct-secheresse-emmanuel-macron-dans-les-hautes-alpes-pour-presenter-un-plan-sur-la-gestion-de-l-eau_5741273.html

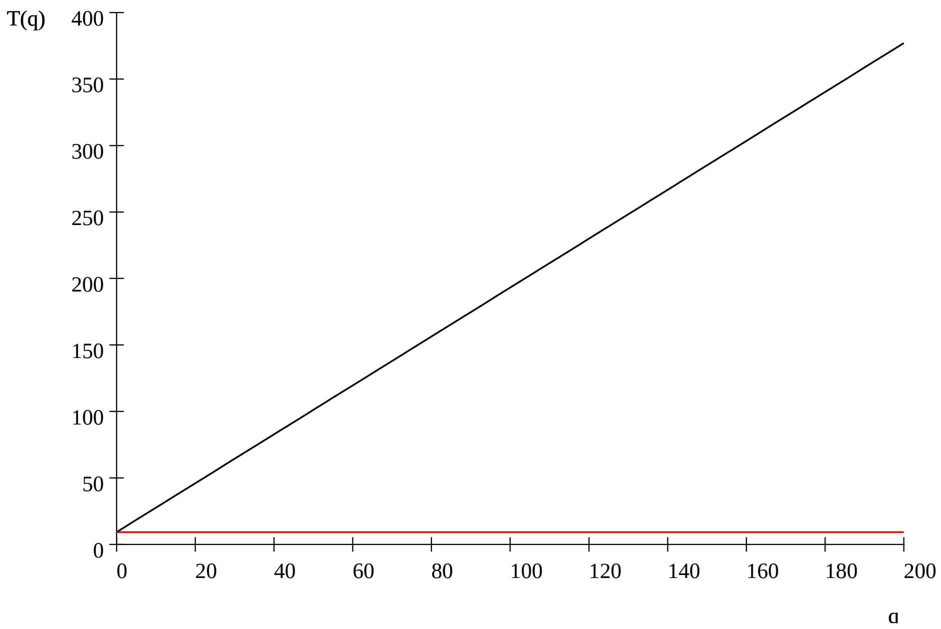


Figure 3 : EP pricing - Melun - Sénart (77 - France) with $F = 9.2629$ *(per trimester) and* $\pi = 1.8389$ *(unit: m³)*

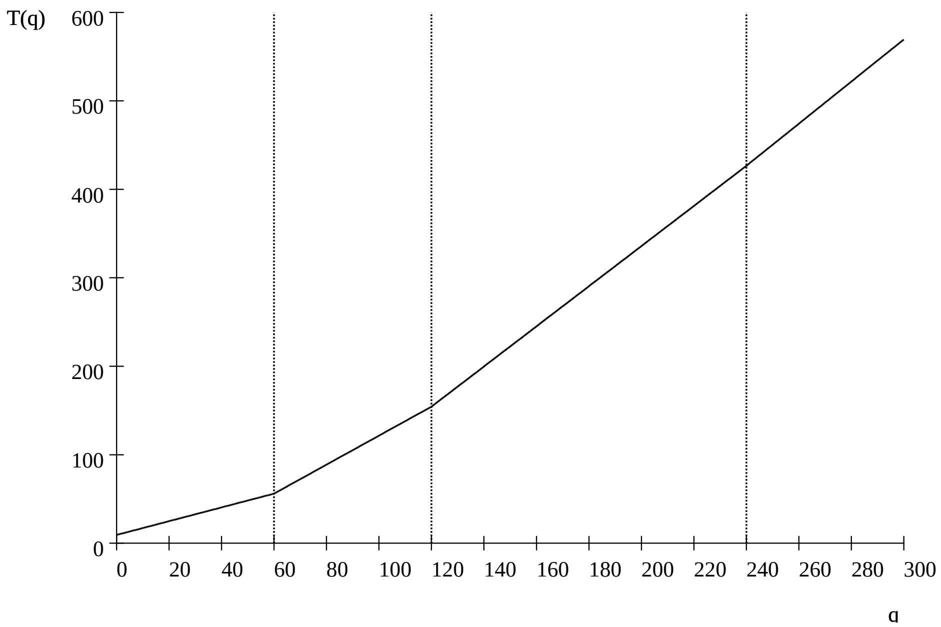


Figure 4 : EP pricing - Saint Paul (974 - France) with $F = 9.38$, $k_1 = 60$, $k_2 = 120$, $k_3 = 240$ *(per trimester) and* $\pi_1 = 0.778$, $\pi_2 = 1.639$, $\pi_3 = 2.268$, $\pi_4 = 2.38$ *(unit: m³).*

4.2 Increasing Block Tariffs (IBT)

4.2.1 Definition

The following notations are used:

- *p* is the number of consumption blocks,
- $I_1,\ I_2,\ I_3\ \ldots$ are the consumption blocks of the form $I_1=[0,k_1],\ I_2=]k_1,k_2],\ I_3=]k_2,k_3]\ \ldots$ with $k_0=0,\ k_1,\ k_2,k_3\ \ldots$ the tariff thresholds,

(with this formulation, a consumption equal to k_1 is considered to be in Block 1, a consumption equal to k_2 is considered to be in Block 2 and so on) and:

- $\pi_1,\ \pi_2,\ \pi_3\ \ldots$ are the unit prices (per m^3) in each consumption block with $\pi_1<\pi_2<\pi_3<\ldots$.

Then, we have:

- in the case of two consumption blocks (denoted as “IBT2” in the following):

$$T(q)=\begin{cases} F+\pi_1 q & if \quad 0\le q\le k_1 \\ F+\pi_1 k_1+\pi_2(q-k_1) & if \quad q>k_1 \end{cases} \tag{4.5}$$

with $\theta_2=(F,\pi_1,\pi_2,k_1)$ the vector of tariff parameters for an IBT2;

- in the case of three consumption blocks (denoted as “IBT3” in the following):

$$T(q)=\begin{cases} F+\pi_1 q & if \quad 0\le q\le k_1 \\ F+\pi_1 k_1+\pi_2(q-k_1) & if \quad k_1<q\le k_2 \\ F+\pi_1 k_1+\pi_2(k_2-k_1)+\pi_3(q-k_2) & if \quad q>k_2 \end{cases} \tag{4.6}$$

with $\theta_3=(F,\pi_1,\pi_2,\pi_3,k_1,k_2)$ the vector of tariff parameters for an IBT3;

- in the case of four consumption blocks (denoted as “IBT4” in the following):

$$T(q)=\begin{cases} F+\pi_1 q & if \quad 0\le q\le k_1 \\ F+\pi_1 k_1+\pi_2(q-k_1) & if \quad k_1<q\le k_2 \\ F+\pi_1 k_1+\pi_2(k_2-k_1)+\pi_3(q-k_2) & if \quad k_2<q\le k_3 \\ F+\pi_1 k_1+\pi_2(k_2-k_1)+\pi_3(k_3-k_2)+\pi_4(q-k_3) & if \quad q>k_3 \end{cases} \tag{4.7}$$

with $\theta_4=(F,\pi_1,\pi_2,\pi_3,\pi_4,k_1,k_2,k_3)$ the vector of tariff parameters for an IBT4

and so on.

4.2.2 Example: EP Tariff for Saint Paul [974]

By way of illustration, Figure 4, page 33, gives the graph of the EP tariff for the city of Saint Paul (Reunion Island) for which $p = 4$ with a billing period that is quarterly and a vector of tariff parameters given by:

$$\theta_4 = (F, \pi_1, \pi_2, \pi_3, \pi_4, k_1, k_2, k_3) = (9.38, 0.778, 1.639, 2.268, 2.38, 60, 120, 240)$$

(this water tariff also corresponds to the one that is preloaded in the MSM). Neglecting VAT and environmental charges (which are introduced later), then we have:

- for a consumption in Block 1 ($0 \le q \le 60$):

$$T(q) = 9.38 + 0.778 \times q$$

- for a consumption in Block 2 ($60 < q \le 120$):

$$T(q) = 9.38 + 0.778 \times 60 + 1.639 \times (q - 60)$$

- for a consumption in Block 3 ($120 < q \le 240$):

$$T(q) = 9.38 + 0.778 \times 60 + 1.639 \times (120 - 60) + 2.268 \times (q - 120)$$

- for a consumption in Block 4 ($q > 240$):

$$T(q) = 9.38 + 0.778 \times 60 + 1.639 \times (120 - 60) + 2.268 \times (240 - 120) + 2.38 \times (q - 240)$$

Mathematically, the Tariff function $T(q)$ is thus a piecewise linear function whose expression is given (ultimately) by:

$$T(q) = \begin{cases} 9.38 + 0.778 \times q & if \quad 0 \le q \le 60 \\ -42.28 + 1.639 \times q & if \quad 60 < q \le 120 \\ -117.76 + 2.268 \times q & if \quad 120 < q \le 240 \\ -144.64 + 2.38 \times q & if \quad q > 240 \end{cases} \qquad (4.8)$$

■

4.2.3 Some Properties

Definition Marginal Price The marginal price is the price of the consumption block in which the household is located with / its drinking water consumption takes place.

Property Apart from thresholds k_1, k_2, k_3 $\ldots$ at which it is not well defined, the marginal price indicates how much the household's water bill varies when the household varies its water consumption, up or down, by one unit (m^3). As such, it also measures the marginal cost of consumption faced by household *i* at consumption $q_i \in I_j$.

Property – Nordin's D In fine, an IBT can be rewritten as:

$$T(q) = F - D(q) + \pi(q) \cdot q \tag{4.9}$$

with $\pi = \pi(q)$ the marginal price function (or marginal price scale) and $D = D(q)$ the Nordin's D function (also called the "Difference variable"), defined as follows:

IBT2

$$\pi(q) = \begin{cases} \pi_1 & if \quad 0 \le q \le k_1 \\ \pi_2 & if \quad k_1 < q \le k_2 \end{cases} \qquad D(q) = \begin{cases} D_1 & if \quad 0 \le q \le k_1 \\ D_2 & if \quad k_1 < q \le k_2 \end{cases}$$

IBT3

$$\pi(q) = \begin{cases} \pi_1 & if \quad 0 \le q \le k_1 \\ \pi_2 & if \quad k_1 < q \le k_2 \\ \pi_3 & if \quad k_2 < q \le k_3 \end{cases} \qquad D(q) = \begin{cases} D_1 & if \quad 0 \le q \le k_1 \\ D_2 & if \quad k_1 < q \le k_2 \\ D_3 & if \quad k_2 < q \le k_3 \end{cases}$$

IBT4

$$\pi(q) = \begin{cases} \pi_1 & if \quad 0 \le q \le k_1 \\ \pi_2 & if \quad k_1 < q \le k_2 \\ \pi_3 & if \quad k_2 < q \le k_3 \\ \pi_4 & if \quad k_3 < q \end{cases} \qquad D(q) = \begin{cases} D_1 & if \quad 0 \le q \le k_1 \\ D_2 & if \quad k_1 < q \le k_2 \\ D_3 & if \quad k_2 < q \le k_3 \\ D_4 & if \quad k_3 < q \end{cases}$$

with:

$$D_1 = 0$$

$$D_2 = (\pi_2 - \pi_1) \cdot k_1,$$

$$D_3 = (\pi_2 - \pi_1) \cdot k_1 + (\pi_3 - \pi_2) \cdot k_2 = D_2 + (\pi_3 - \pi_2) \cdot k_2,$$

$$D_4 = (\pi_2 - \pi_1) \cdot k_1 + (\pi_3 - \pi_2) \cdot k_2 + (\pi_4 - \pi_3) \cdot k_3 = D_3 + (\pi_4 - \pi_3) \cdot k_3,$$

...

the values for block 1, block 2, block 3, block 4 ... of Nordin's D.

Interpretation of Nordin's D (or "Virtual Reimbursement") Nordin's D formally corresponds to a difference between two levels of expenditure:

- the amount of the bill actually paid by the household $T(q)=F-D_j+\pi_j q$ for a water consumption q located in block j;
- the amount that would have been charged for the same level of consumption $q\in I_j$ if each of the q units had been sold at the marginal price π_j,

i.e. if a two-part tariff with parameters (F,π_j) had been applied (and then $T_2(q)=F+\pi_j q$). In this way, the calculation of Nordin's D measures the extent of the savings (or monetary gains) made by the household and generated by the price progressivity of the tariff. As such, Nordin's D can also be seen as an indicator measuring the degree of tariff progressivity.

Illustration Taking the elements relating to the calculation of the Billing function of the EP tariff for the municipality of Saint Paul, we have:

- For a consumption in Block 2 ($60<q\le 120$):

$$\begin{aligned}
T(q) &= F+\pi_1 k_1+\pi_2(q-k_1)\\
&= 9.38+0.778\times 60+1.639\times(q-60)\\
&= 9.38-(1.639-0.778)\times 60+1.639\times q\\
&= 9.38-0.861\times 60+1.639\times q\\
&= 9.38-51.66+1.639\times q
\end{aligned}$$

- For a consumption in Block 3 ($120<q\le 240$):

$$\begin{aligned}
T(q) &= F+\pi_1 k_1+\pi_2(k_2-k_1)+\pi_3(q-k_2)\\
&= 9.38+0.778\times 60+1.639\times(120-60)+2.268\times(q-120)\\
&= 9.38-(1.639-0.778)\times 60-(2.268-1.639)\times 120+2.268\times q\\
&= 9.38-0.861\times 60-0.629\times 120+2.268\times q\\
&= 9.38-51.66-75.48+2.268\times q = 9.38-127.14+2.268\times q
\end{aligned}$$

- For a consumption in Block 4 ($240<q$):

$$\begin{aligned}
T(q) &= F+\pi_1 k_1+\pi_2(k_2-k_1)+\pi_3(k_3-k_2)+\pi_4(q-k_3)\\
&= 9.38+0.778\times 60+1.639\times(120-60)+2.268\times(240-120)+2.38\times(q-240)\\
&= 9.38-(1.639-0.778)\times 60-(2.268-1.639)\times 120-(2.38-2.268)\times 240+2.38\times q\\
&= 9.38-0.861\times 60-0.629\times 120-0.112\times 240+2.38\times q\\
&= 9.38-51.66-75.48-26.88+2.38\times q\\
&= 9.38-154.02+2.38\times q
\end{aligned}$$

Mathematically, the Tariff function given by the equation (4.8) on page 35 can then be rewritten as :

$T(q) = 9.38 - D(q) + \pi(q)q$

with:

$$\pi = \pi(q) = \begin{cases} 0.778 & if \quad 0 \leq q \leq 60 \\ 1.639 & if \quad 60 < q \leq 120 \\ 2.268 & if \quad 120 < q \leq 240 \\ 2.38 & if \quad 240 < q \end{cases}$$

the graph of the unit price scale (of the EP Tariff of the city of Saint Paul) and

$$D = D(q) = \begin{cases} 0 & if \quad 0 \leq q \leq 60 \\ 51.66 & if \quad 60 < q \leq 120 \\ 127.14 & if \quad 120 < q \leq 240 \\ 154.02 & if \quad 240 < q \end{cases}$$

the graph of the Nordin D function / Virtual reimbursement (of the EP Tariff of the Commune of Saint Paul). See Figure 5 & Figure 6, for an illustration of the equivalence; and Figure 7 & Figure 8, page 39, for a representation of the unit price scale and of the graph of Nordin's D function.

(2) **On tariff progressiveness** By analogy with Taxation theory (see Lambert [2001]), an EP / EPA tariff is said to be structurally progressive (respectively structurally degressive) when the average price (in fact, the household average cost of consumption):

$$\bar{T}(q) = \frac{T(q)}{q} = \ldots \tag{4.10}$$

is increasing (respectively is decreasing) with the level of consumption q (this notion of structural progressivity / degressivity can be local). In this case, the following properties should be borne in mind.

- A two-part tariff (F, π) for which:

$$\bar{T}(q) = \frac{T(q)}{q} = \frac{F + \pi q}{q} = \frac{F}{q} + \pi \tag{4.11}$$

is everywhere (i.e. at any point) structurally degressive (as soon as $F > 0$).

- The only scheme that is neither structurally progressive nor structurally degressive is linear pricing $T(q) = \pi q$ (with, in this case, $\bar{T}(q) = \pi = c^{ste}$).

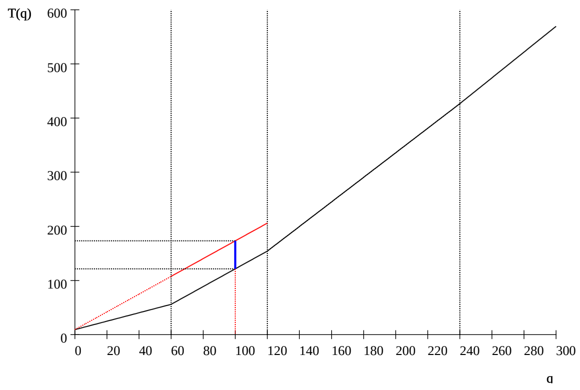


Figure 5 : Nordin D – Block 2

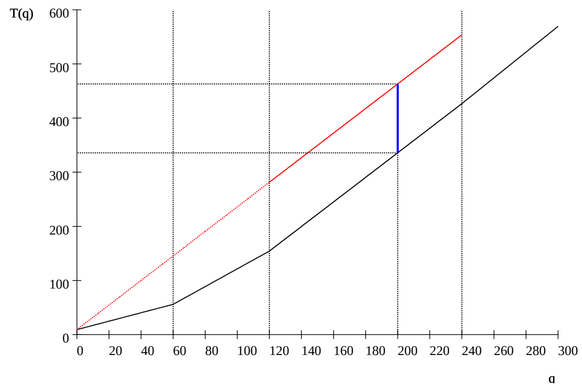


Figure 6 : Nordin D – Block 3

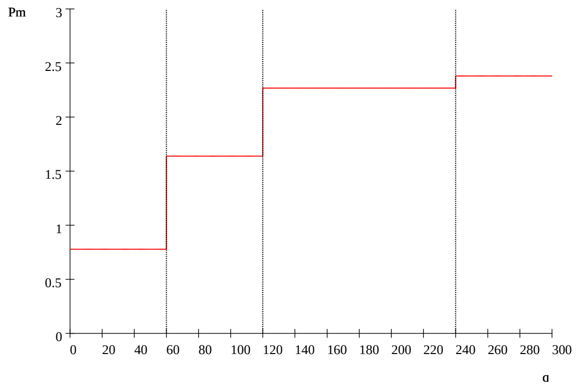


Figure 7 : Marginal price graph $\pi = \pi(q)$

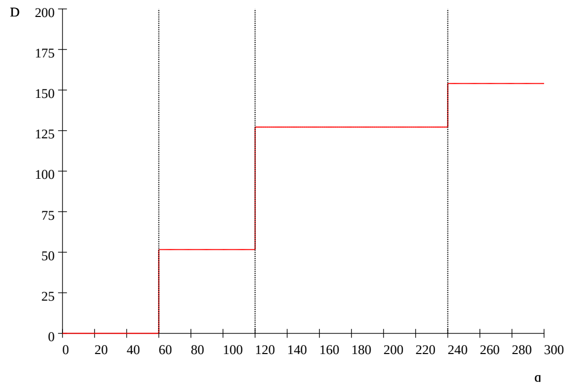


Figure 8 : Nordin D graph $D = D(q)$

- An IBT is structurally degressive for sure in block 1 with:

$$\bar{T}(q)=\frac{T(q)}{q}=\frac{F+\pi_1 q}{q}=\frac{F}{q}+\pi_1 \quad \text{for} \quad q\in I_1=[0,k_1] \tag{4.12}$$

(this property also shows that an IBT is locally equivalent to a Two-part tariff in block 1).

- An IBT may be structurally degressive everywhere (i.e. at any point).

- An IBT becomes structurally progressive from the tariff threshold $\bar{k}$ where it exists, above which Nordin's D exceeds the subscription amount F .

Details Faced with increasing block tariff, the average cost of a consumption $q\in I_j$ is given by:

$$\bar{T}(q)=\frac{T(q)}{q}=\frac{F-D_j+\pi_j q}{q}=\frac{F-D_j}{q}+\pi_j \tag{4.13}$$

with:

$$\frac{\partial\bar{T}}{\partial q}=-\frac{F-D_j}{q^2}\geq 0\Leftrightarrow D_j\geq F \tag{4.14}$$

Accordingly, the tariff will be structurally progressive (respectively structurally degressive) in the consumption blocks for which Nordin's D (virtual reimbursement) is higher (respectively lower) than the amount of the subscription fee F . The values D_j being increasing in j with:

$D_j=D_{j-1}+(\pi_j-\pi_{j-1})k_{j-1}$ and $\pi_j>\pi_{j-1}$,

an IBT that would be structurally progressive in a block j is also progressive in a higher block and it suffices to locate the tariff threshold, denoted $\bar{k}$, beyond which Nordin's D becomes greater than the amount of the subscription fee to locate the consumption intervals in which the tariff $T(\cdot)$ is successively structurally degressive, then structurally progressive. As pointed out, this threshold $\bar{k}$ may not exist, i.e. the tariff parameters may be such that $D_j<F$ for any *j*, and in this case an IBT is everywhere structurally degressive.

Example As a reminder, the EP tariff for the city of Saint-Paul is an IBT4 with $F=9.38$, $k_1=60$, $k_2=120$, $k_3=240$, and $\pi_1=0.778$, $\pi_2=1.639$, $\pi_3=2.268$ and $\pi_4=2.38$. The one of the commune of Saint-Denis (Reunion Island - France) is an IBT3 with:

$$\theta_3=(F,\pi_1,\pi_2,\pi_3,k_1,k_2)=(7.9944,0.78198,0.8655,1.1494,45,90)$$

See Figure 9 & Figure 10, page 41. Water pricing in Saint Paul is structurally progressive from block 2 with a $\bar{k}$ equal to 60 m³, while pricing in Saint-Denis is structurally progressive from block 3 with a $\bar{k}$ equal to 90 m³. Saint-Paul's pricing system therefore becomes (structurally) progressive more quickly than Saint-Denis's one.

Note Calculating the degree of structural progressivity of an IBT is a complex issue (Suarez-Varela & al. [2015]). In practice, the calculation of this threshold $\bar{k}$ (above which an IBT is structurally progressive) is often carried out by stakeholders to assess the incentive nature of the tariff (see for instance Office de l'Eau - Réunion [2019]).

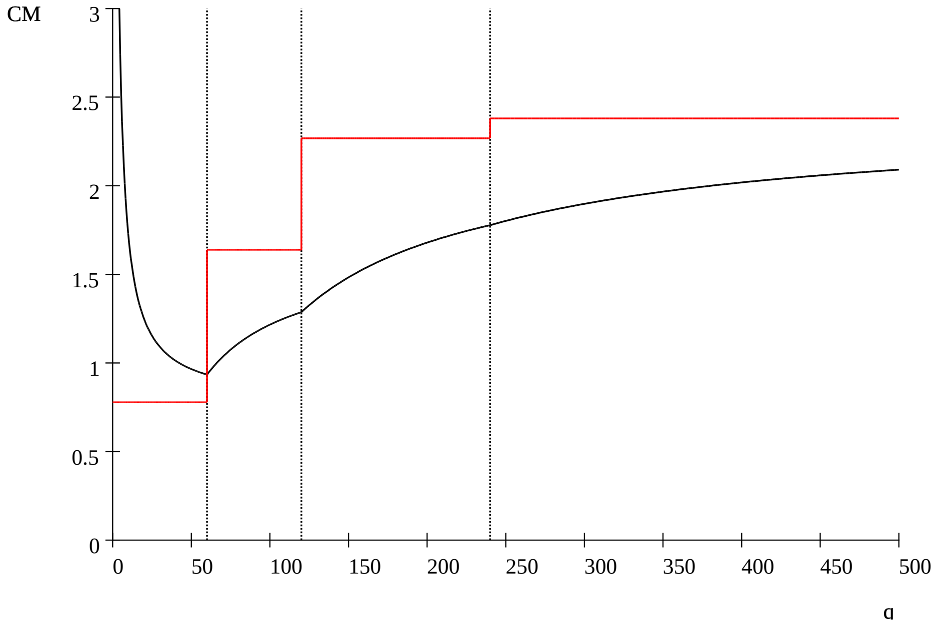


Figure 9 : Structural progressivity - EP *tariff 2018 Saint Paul*

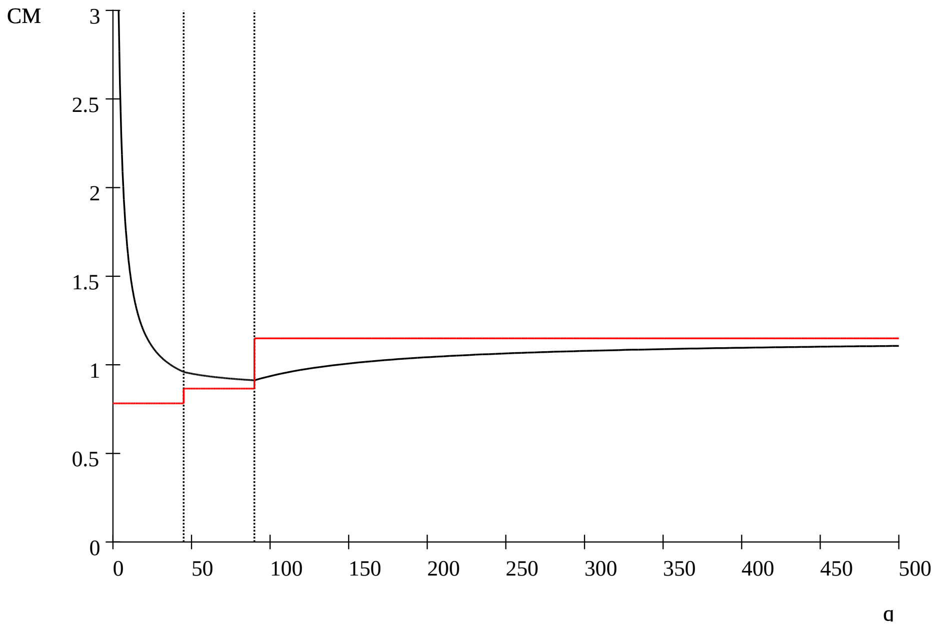


Figure 10 : Structural progression - EP *tariff 2018 Saint Denis*

(3) In what follows, a clear distinction has to be made between the average cost (or average price) of household consumption, and therefore the average function:

$$\bar{T}(q) = \frac{T(q)}{q} = \begin{cases} \dfrac{F}{q} + \pi_1 & if \quad 0 \le q \le k_1 \\ \dfrac{F - D_2}{q} + \pi_2 & if \quad k_1 < q \le k_2 \\ \dfrac{F - D_3}{q} + \pi_3 & if \quad k_2 < q \le k_3 \\ \ldots \end{cases} \quad (4.15)$$

with $D_2 = (\pi_2 - \pi_1)k_1$, $D_3 = D_2 + (\pi_3 - \pi_2)k_2 \ldots$ the values of Nordin's D in blocks 2, 3..., and the marginal cost of consumption faced by the Household. Mathematically, the latter is defined as the derivative of the Tariff function with respect to q:

$$T'(q) = \begin{cases} \pi_1 & if \quad 0 \le q < k_1 \\ \pi_2 & if \quad k_1 < q < k_2 \\ \pi_3 & if \quad k_2 < q < k_3 \\ \ldots \end{cases} \quad (4.16)$$

and corresponds geometrically to the graph of the unit price scale $\pi(q)$ (neglecting the problems of non-derivability at $q = k_1$, $q = k_2 \ldots$). It should be noted that this is this unit price schedule (and not the average cost of consumption) that the household has to consider for an optimal management of its domestic water uses (see paragraph 5.2.3 "Tariff perception").

(4) **On Terminology** For the main, a tariff is a function $T(\cdot)$ that specifies, for each level of water consumption q, the amount $T(q)$ that the household will have to pay if it intends to acquire those q units of service. Neglecting the non-derivability problems that arise with an IBT (see above), the derivative function $T'(q)$ of the tariff function is also called the marginal price curve and is sometimes noted $\pi(q)$ in the literature (to emphasise the fact that the household faces a price schedule). At a given level of consumption $q = q_0$, the value taken by this function $\pi(q_0) = T'(q_0)$ is interpreted as giving the monetary cost borne by the consumer if he increases his consumption by one unit and, also, the monetary gain realised by the consumer if he reduces his consumption by one unit. For this reason, one speaks of the price of the q-th unit of consumption or, more directly, of the price of the q-th increment (Wilson [1993]). The latter is also known as the household marginal cost of consumption.

In this context, non-linear pricing can be seen as the pricing of a product line made up of q successive units of the same generic product, in this case the EP or EPA supply, which are then treated as different goods sold at (potentially) different prices. However, the purchase of a specific unit requires the purchase of the previous units in the sequence, i.e. the purchase of the second unit requires the purchase of the first, and so on. Within this framework, a tariff is said to be progressive (respectively degressive) if the unit price scale / the marginal price curve / the marginal cost of consumption $\pi(q) = T'(q)$ is globally increasing (respectively globally decreasing) in q. If the marginal cost of consumption is now constant in all q, the tariff is said to be flat or uniform. Flat rates include linear tariffs $T(q) = \pi q$ and, also, two-part tariffs where $T(q) = F + \pi q$, with $F > 0$ the access fee.

4.3 Miscellaneous - getting started

4.3.1 The tabs "EP and A" tariffs

The user is asked to enter the characteristics of the tariff, that is the amount of the subscription F, the sizes of the *p* consumption blocks with the setting of the thresholds k_1, k_2, ..., k_{p-1} and the unit prices (per cubic metre) for each of the consumption blocks π_1, π_2, ..., π_p, for, successively, (1) the drinking water supply (EP service) and (2) the (collective) wastewater supply (A service). These amounts to be filled in must be given exclusive of taxes and environmental charges (which are introduced later). The billing period (unit of time) is that of a consumption quarter (in the case of half-yearly billing, the user must enter a subscription amount divided by 2 and the sizes of the consumption blocks also divided by 2; the prices per cubic metre do not need to be adjusted). By default, the tool displays the progressive tariff for the City of Saint Paul (preloaded) with an “EP” tariff (water utility) given by:

$$\theta_4^{\text{EP}} = (F, \pi_1, \pi_2, \pi_3, \pi_4, k_1, k_2, k_3) = (9.38, 0.778, 1.639, 2.268, 2.38, 60, 120, 240)$$

and an “A” tariff (wastewater utility) given by:

$$\theta_4^{\text{A}} = (F, \pi_1, \pi_2, \pi_3, \pi_4, k_1, k_2, k_3) = (10.49, 1.3, 2.12, 2.21, 2.30, 60, 120, 240)$$

Based on this information entered by the user, the tool then consolidates the two tariffs into an EPA tariff (drinking water and wastewater utility). The latter in the case of the City of Saint Paul writes simply as:

$$\theta_4^{\text{EPA}} = (F, \pi_1, \pi_2, \pi_3, \pi_4, k_1, k_2, k_3) = (19.87, 2.078, 3.759, 4.478, 4.68, 60, 120, 240)$$

It is to note that this consolidation operation (merging the 2 tariffs) is also carried out when the number of consumption blocks and/or the sizes of the consumption blocks for the drinking water supply and the wastewater supply differ. Besides, there is no limit on the number of blocks the user wishes to set, and it is sufficient to enter the price of a single consumption block to enter a two-part tariff (F, π) (while deleting the 3 higher blocks entered by default) or a linear tariff (in the latter case, the subscription amount for the drinking water service and/or the wastewater service must be set to 0). It should also be noted that it is not possible for the user to enter a degressive tariff for which the econometric demand model estimated by BCP [2014] (used in the Demand Module) does not apply.

To conclude, two points should be kept in mind. First, as the tool is based on French regulations governing the pricing of drinking water and wastewater services, a household that is not connected to the sewerage network is not billed for collective sanitation. Without adjustment of the sample, this is initially the case for just over half of the households (54.1%) in the data file describing the Population to which the tariff evaluated/tested by the user is designed to apply (in the Household database, there is a dummy variable indicating whether or not the household is connected to the sewerage network, and therefore whether it is charged the EPA tariff or the EP tariff only). Secund, the calibration if the IBT is implemented by the user on the basis of 4 pieces of information provided by the tool.

The first relates to the distribution of volumes of water to cover the basic needs within the population of the households. This information, which is necessary to set with legibility the sizes of the first consumption blocks (which are subsidised) and take the measure of exclusion errors (linked to the fact that part of the population will see part of its basic consumption unsubsidised, or even taxed/marged), is provided with the display of the Pen's Parade of basic consumption and, on the same diagram, the related concentration curve of the mass of basic consumption (see figure 2.2, page 29, as well as the information given in footnote **Erreur ! Signet non défini.**, page 27).

The second piece of information relates to the affordability of TBSE for the general EP / EPA service, which is the reference tariff. The infographic in question is a scatter plot with :

- on the x-axis, the standard of living of Households (from the Population Module) calculated as household income (net of income tax) divided by an equivalent number of adults (consumption units) for the modified OECD scale;
- on the y-axis, the Potential Affordability Ratio (PAR), which gives the proportion of income that a Household must devote to meeting its basic needs with the "TBSE" reference tariff (see section 3.1).

This scatter plot is then completed:

- of the (vertical) poverty line, i.e. the threshold standard of living (entered beforehand by the user) below which the Household is considered to be poor,
- of the (horizontal) water poverty line, i.e. the threshold value for the PAR (previously entered by the user) above which the Household is considered to face an affordability issue,

so as to identify 4 groups within the population depending on whether the Household (*i*) is a poor one or not and (*ii*) initially faces an affordability issue or not. The other two infographics provide information on (i) the level of TBSE consumption and its correlation with the standard of living of Households (this last characteristic relating to the environment of the decision problem facing the user is a key parameter for the potentially redistributive or antiredistributive nature of the system of cross-subsidies that is implemented by a IBT of the social incentive type) with the display of the corresponding scatterplot and (ii) the TBSE deviations from the first-best social optimum (which makes it possible to assess the potentially high level of TBSE consumption). Finally, the user is reminded of the values of the price and income elasticities of the demand functions (which will determine how the sales revenue of the water company will change with the switch from TBSE to IBT tested/evaluated by the user).

4.3.2 Taxation

The price of water and wastewater services is also affected by taxation, with a value-added tax (VAT) and ecological taxes (excise duties).

VAT (Value Added Tax) is an indirect tax (collected by the operator and paid to the State) which relates to the value of the goods (ad valorem tax). The amount of VAT (which is a tax borne by the household) is calculated by applying a rate to the price of "EP" service, on the one hand, and to the price of "A" service, on the other.

Initially, the French tax system defined 4 VAT rates with a standard rate of 20%, some reduced rates of 5.5% and 10%, and a special rate of 2.1%. In mainland France, the reduced rate of 5.5% (like gas and electricity) is applied to public drinking water supply. Reunion Island, which is an overseas department, applies a lower rate than the national rate, corresponding to the special rate of 2.10%. The VAT rate for sanitation is 10% for mainland France and 2.10% for the French Overseas Territories. Another special feature of French law is that the (collective) sanitation service can be exempted from VAT when the service is provided by a public firm.

Given these various institutional constraints, the user has the option of defining 2 rates:

- a VAT rate for the public drinking water utility;
- a VAT rate (which may be set to 0) for the public sewerage utility.

This information should be entered in the VAT field in the Taxation section of the General Data tab of the microsimulation model.

Environmental charges Drinking water consumption and wastewater services are also subject to taxes in the form of excise duties, i.e. a sum in euros per unit consumed. These taxes are collected by the operator for the benefit of the local Water Agency, which uses them to finance various actions (such as investment aid to protect water resources).

These taxes, noted r_{EP} and r_{A}, the amounts of which are entered by the user, are considered to be ecological taxes. A special feature of French law is that these charges are also subject to VAT. This information should be entered in the Duties field in the Taxation section of the General Data tab of the microsimulation model.

Combined with the tariff data describing the pricing for water and wastewater services (which are entered by the user), the combination of VAT and environmental duties leads to the calculation of prices, including VAT, for the public drinking water service and the public wastewater service, one per block, as summarised in Table 2 on next page (it should be noted that, given the stepwise increase in service prices, the VAT mechanism makes that the amount of tax per unit of service also increases in steps with the level of consumption, i.e. an Increasing Block Tariff generates an Increasing Block unit Tax Fee). These price scales are then used to calculate household water consumption in the Demand module (making use of the econometric model of household water demand) presented in section 4.

Service			Price excl. VAT	Excice duty	VAT	Taxes	Price incl. VAT
Drinking water					$t_{\mathrm{EP}} = 2.10\%$		
Subscription fee			F_{EP}		t_{EP}	$t_{\mathrm{EP}}F_{\mathrm{EP}}$	$(1+t_{\mathrm{EP}})F_{\mathrm{EP}}$
T1-EP	0	k_1	π_1^{EP}	r_{EP}	t_{EP}	$t_{\mathrm{EP}}(\pi_1^{\mathrm{EP}}+r_{\mathrm{EP}})$	$(1+t_{\mathrm{EP}})(\pi_1^{\mathrm{EP}}+r_{\mathrm{EP}})$
T2-EP	k_1	k_2	π_2^{EP}	r_{EP}	t_{EP}	$t_{\mathrm{EP}}(\pi_2^{\mathrm{EP}}+r_{\mathrm{EP}})$	$(1+t_{\mathrm{EP}})(\pi_2^{\mathrm{EP}}+r_{\mathrm{EP}})$
T3-EP	k_2	k_3	π_3^{EP}	r_{EP}	t_{EP}	$t_{\mathrm{EP}}(\pi_3^{\mathrm{EP}}+r_{\mathrm{EP}})$	$(1+t_{\mathrm{EP}})(\pi_3^{\mathrm{EP}}+r_{\mathrm{EP}})$
T4-EP	k_3	$+\infty$	π_4^{EP}	r_{EP}	t_{EP}	$t_{\mathrm{EP}}(\pi_4^{\mathrm{EP}}+r_{\mathrm{EP}})$	$(1+t_{\mathrm{EP}})(\pi_4^{\mathrm{EP}}+r_{\mathrm{EP}})$
…							
Wastewater					$t_{\mathrm{A}} = 10\%$		
Subscription fee			F_{A}		t_{A}	$t_{\mathrm{A}}F_{\mathrm{A}}$	$(1+t_{\mathrm{A}})F_{\mathrm{A}}$
T1-A	0	k_1	π_1^{A}	r_{A}	t_{A}	$t_{\mathrm{A}}(\pi_1^{\mathrm{A}}+r_{\mathrm{A}})$	$(1+t_{\mathrm{A}})(\pi_1^{\mathrm{A}}+r_{\mathrm{A}})$
T2-A	k_1	k_2	π_2^{A}	r_{A}	t_{A}	$t_{\mathrm{A}}(\pi_2^{\mathrm{A}}+r_{\mathrm{A}})$	$(1+t_{\mathrm{A}})(\pi_2^{\mathrm{A}}+r_{\mathrm{A}})$
T3-A	k_2	k_3	π_3^{A}	r_{A}	t_{A}	$t_{\mathrm{A}}(\pi_3^{\mathrm{A}}+r_{\mathrm{A}})$	$(1+t_{\mathrm{A}})(\pi_3^{\mathrm{A}}+r_{\mathrm{A}})$
T4-A	k_3	$+\infty$	π_4^{A}	r_{A}	t_{A}	$t_{\mathrm{A}}(\pi_4^{\mathrm{A}}+r_{\mathrm{A}})$	$(1+t_{\mathrm{A}})(\pi_4^{\mathrm{A}}+r_{\mathrm{A}})$
…							

Table 2 : Prices (incl. VAT) for drinking water and wastewater services (summary of price scale)

4.3.3 Cost modelling

4.3.3.1 Service production and distribution costs

Basically, service cost is modelled simply by assuming a linear cost function of the form:

$$C(Q) = CF + c \times Q \tag{4.17}$$

where Q is the overall production/service level, CF is the amount of fixed costs and c is the unit variable cost of production, assumed to be constant. This cost parameter also measures the marginal cost of production $C'(Q) = c$, that is the increase in production cost that is borne by the operator when overall production is increased by one unit (1 cubic metre). The latter differs from the unit total cost:

$$\frac{C(Q)}{Q} = \frac{F}{Q} + c > \mathrm{CVM}(Q) = \frac{C(Q) - CF}{Q} = C'(Q) = c \tag{4.18}$$

as soon as $F > 0$. As a general rule, it is estimated that the first component (average fixed cost) represents between 60 and 80% of the cost of the service. Overall production is calculated as the sum of household water consumption:

$$Q = \sum_{i=1}^{n} q_i \tag{4.19}$$

with n the number of domestic subscribers.

The user is asked to enter the values of these two cost parameters F and c for:

- the EP service with $F_{\mathrm{EP}} = \ldots$ and $c_{\mathrm{EP}} = \ldots$,
- the A service with $F_{\mathrm{A}} = \ldots$ and $c_{\mathrm{A}} = \ldots$,

as well as the number of subscribers $n = n_{\mathrm{EP}}$ and n_{A} (this information is entered in the EPA service costs section of the General Data tab). Next, the cost of the general service $C = C_{\mathrm{EP}} + C_{\mathrm{A}}$ that is borne by the operator is calculated as follows:

$$\begin{aligned} C &= C_{\mathrm{EP}} + C_{\mathrm{A}} = C(Q_{\mathrm{EP}}) + C(Q_{\mathrm{A}}) = CF_{\mathrm{EP}} + c_{\mathrm{EP}} \times Q_{\mathrm{EP}} + CF_{\mathrm{A}} + c_{\mathrm{A}} \times Q_{\mathrm{A}} \\ &= CF_{\mathrm{EP}} + CF_{\mathrm{A}} + c_{\mathrm{EP}} \sum_{i=1}^{n} q_i + c_{\mathrm{A}} \sum_{i=1}^{n_{\mathrm{A}}} q_i = CF_{\mathrm{EP}} + CF_{\mathrm{A}} + c_{\mathrm{EP}} \sum_{i=1}^{n-n_{\mathrm{A}}} q_i + (c_{\mathrm{EP}} + c_{\mathrm{A}}) \sum_{i=1}^{n_{\mathrm{A}}} q_i \\ &= CF_{\mathrm{EP}} + CF_{\mathrm{A}} + c_{\mathrm{EP}} \times n_1 \overline{q}_1 + c_{\mathrm{EPA}} \times n_2 \overline{q}_2 \end{aligned} \tag{4.20}$$

where:

- $c_{\mathrm{EPA}} = c_{\mathrm{EP}} + c_{\mathrm{A}}$ is the unit variable cost for the EPA service, that is for the production of one cubic metre of treated drinking water,

- $\overline{q}_1$ is the average water consumption of the $n - n_{\mathrm{A}} = n_1$ households who are not connected to the collective sewerage system :

$$\bar{q}_1 = \frac{1}{n - n_A} \times \sum_{i=1}^{n-n_A} q_i$$

and therefore pay for the drinking water service only;

- $\bar{q}_2$ is the average water consumption of the $n_2 = n_A$ households who are connected to the collective sewerage network :

$$\bar{q}_2 = \frac{1}{n_A} \sum_{i=1}^{n_A} q_i = \frac{1}{n_2} \sum_{i=1}^{n_2} q_i = \bar{q}_{EPA}$$

and who therefore pay for the drinking water and wastewater services.

These two groups of household subscribers (on which a certain number of cross-tabulations are carried out subsequently) are referred in the following as Group 1 and Group 2. The average consumption of these two groups of subscribers, $\bar{q}_1$ and $\bar{q}_2$, is estimated from the averages of the samples of households in Group 1 and Group 2 (which are considered to be representative) in the Population module. On this basis, the (estimated) level of service Q_{EP} is simply calculated as :

$$Q_{EP} = n_1 \times \bar{q}_1 + n_2 \times \bar{q}_2 = (n - n_A) \times \bar{q}_1 + n_A \times \bar{q}_2$$

and the (estimated) level of "A" service as :

$$Q_A = n_2 \times \bar{q}_2 = n_A \times \bar{q}_2$$

given the values $n_{EP} = n$ and n_A that have been entered for the numbers of domestic subscribers by the user. The calculation of domestic consumptions for the "EP" tariff and the "EPA" tariff that are evaluated/tested by the user is based on the econometric demand model in the Demand module. This information on production costs, which is entered by the user, also sets the parameters for the reference tariffs, the TBSE for the EP service alone and the TBSE for the EPA service, from which the gains and losses of the IBT that is tested/evaluated by the user are taken. They also set the cost of the service that household *i* passes on to the operator through its consumption q_i. Assuming an egalitarian distribution of fixed costs over the household subscriber population, the latter can be calculated simply as follows :

$$C_i^{EP}(q_i) = \frac{CF_{EP}}{n_{EP}} + c_{EP} \times q_i$$

$$C_i^{A}(q_i) = \frac{CF_A}{n_A} + c_A \times q_i$$

for the EP service and the A service respectively and :

$$C_i^{EPA}(q_i) = \frac{CF_{EP}}{n_{EP}} + \frac{CF_A}{n_A} + (c_{EP} + c_A) \times q_i$$

for the EPA service, when the household is connected to the collective sewerage network.

4.3.3.2 Environmental cost

As cost recovery must be understood in the full sense of the term, i.e. including the environmental cost, the user is also asked to enter the value of the environmental cost. The model used is precisely as follows.

(1) Based on the principle "water has to return clean to nature", it is assumed that the use by the household of a cubic metre of drinking water pollutes this unit of service to the tune of c_e euros, that is the level of the environmental cost entered by the user.

As usual, the latter is defined as the net cost of the actions that need to be implemented for total depollution of a service unit. On this basis, 2 types of water pollutants are distinguished:

(2) those that are cleaned up through the collective sanitation facility with a residual (not recovered) environmental cost that sets to:

$$\max\left[c_e - c_{\mathrm{A}} - r_{\mathrm{EP}} - r_{\mathrm{A}}, 0\right] = \begin{cases} c_e - c_{\mathrm{A}} - r_{\mathrm{EP}} - r_{\mathrm{A}} & \text{if} \quad c_{\mathrm{A}} + r_{\mathrm{EP}} + r_{\mathrm{A}} \leq c_e \\ 0 & \text{if} \quad c_{\mathrm{A}} + r_{\mathrm{EP}} + r_{\mathrm{A}} = c_e \end{cases} \tag{4.21}$$

(in the latter case, full cost recovery is achieved);

(3) those that are not cleaned up with the water consumptions of non-connected household to the public wastewater collection system and a residual (not recovered) environmental cost that sets to:

$$\max\left[c_e - r_{\mathrm{EP}}, 0\right] = c_e - r_{\mathrm{EP}} \tag{4.22}$$

Within this framework, it is therefore considered that non-collective wastewater facilities are fully inefficient concerning the cleaning activities. This strong assumption is nevertheless supported by the fact that a large number of individual systems do not appear to be up to standard in Reunion Island (according to several professionals) and, also, by the much lower user cost of individual systems, compared to the amount of bills paid by residential consumers for the public wastewater collection system (what suggests they are of poorer purification quality). It should also be emphasized that water-related amenities (which have an impact on the management rule for the implementation of an efficient (Pareto optimal) allocation of the water resources) are not considered in the model.

4.3.3.3 On subsidies (and "taxation")

Given the principle "water pays for water", a progressive pricing system of the social incentive type, which aims to subsidise households for their basic needs, implements some subsidies and "taxes" (in fact, margins that contribute to service funding) that the tool enables to measure. To do so, one makes use of the information provided by the user and that relates to the amount of fixed production costs, on the one hand, and the value of the unit variable production cost, assumed to be constant, on the other hand, for the drinking water utility and the wastewater utility respectively. This information then sets the parameters of the TBSE, which is the reference pricing, with an access fee that is billed by the operator at cost price with $F = F_{\mathrm{TBSE}} = CF/n$, and a service unit that is also billed by the operator at cost price, with $\pi = \pi_{\mathrm{TBSE}} = c$ (It should be noted that this calculation is performed for each service, "EP" vs. "A", separately). On this basis, the tool distinguishes between 2 types of subsidy/taxation.

The first relates to the subsidy/taxation of the access fee with a contribution to service funding given by:

$$C_{i0} = F_{\text{TBSE}} - F = \frac{CF}{n} - F \tag{4.23}$$

(a negative value for C_{i0} indicates a subsidy on the access fee, a positive value a "tax").

The second relates to (gross and net) subsidies and (gross and net) taxes on household water consumption that are implemented by the tariff.

Figure 11 illustrates the latter for the case of the Saint Paul EP pricing. The (marginal) price schedule $\pi = \pi(q)$ is shown there in red. The green line shows the marginal cost / unit variable cost (of the EP service) c, which has been set (arbitrarily) at €1.5 (a rather high value a priori). Then, as is apparent, in the case under consideration here:

- the first 60 units / units from 0 to 60 m³ are subsidised at a rate of $\sigma_1 = 1.5 - 0.778 = 0.722€$ per m^3,
- the next 60 units / units of 60 to 120 m³ are "taxed" at a rate of $\tau_2 = 1.639 - 1.5 = 0.139€$ per m^3,
- a "tax" is levied on block 3 units, amounting to $\tau_3 = 2.268 - 1.5 = 0.768€$ per m^3, and on block 4 units, amounting to $\tau_4 = 2.38 - 1.5 = 0.88€$ per m^3.

The Table 3, page 52, then gives the values of the various (gross and net) subsidies on the consumption and per subscriber, Access Fee Included (what is denoted as “DAI”, for “**D**roit d’**A**ccès **I**nclus” in French), for different levels of consumption relating to the distribution of water consumption in this municipality (a negative value indicates a subsidy). The point is that, for the (here high) value of the (presumed) unit variable cost c, the financing of the service is based on a fringe of large consumers (less than 10% of subscribers). The tool then measures these subsidies and taxes for each household in the Population module, based on consumption q_i and basic consumption $\underline{q}_i$ calculated from the demand function in the Demand module, for the tariff that is evaluated/tested by the user. These different amounts are collected in the Invoices module (see section 5) and used to calculate some indicators for measuring the Incentive Effect of pricing, Equity and Quality of the Funding that are displayed in the Evaluation Module[10].

Remark It should also be emphasized that, given the mechanism of VAT, the implementation of a progressive pricing system of the social incentive type also implies a system of subsidies/taxations on the part of the State ("it is as if") on the access fee and consumption of block 1, block 2 etc. For this reason, a distinction is made for the purposes of analysis:

---

[10] To list them, the tool calculates (for each household in the Population module) gross subsidies and gross taxes on the Right of Access, basic consumption (i.e. units of service meeting basic uses), the captive part of demand not meeting basic uses (in the BCP demand model [2014], the part of consumption meeting water uses for garden maintenance and swimming pool maintenance), and overconsumption (linked to poor tariff perception) which is generated by the EP tariff, the A tariff and, consequently, the EPA tariff. On this basis, these various amounts are aggregated to calculate gross and net subsidies on the consumption of the Subscriber (Access Fee Excluded) and per Subscriber (Access Fee Included). These total subsidies and taxes are also broken down into subsidies and taxes on the Basic Service (Access Fee Inlcuded), on basic consumption (Access Fee Excluded) and on non-basic consumption (these aggregated flows are then necessarily Access Fee Excluded).

(i) the subsidy/taxation system implemented by the operator with the application of the IBT,

(ii) the system of subsidies/taxations implicitly implemented by the State, taking into account the IBT applied by the operator, with the VAT mechanism

(iii) the system of subsidies/taxes including VAT that households ultimately face.

These subsidy/taxation systems must also be differentiated according to services, with (i) the system implemented for the drinking water service, (ii) the one implemented for wastewater service and, as a result, (iii) the one implemented for the joined drinking water and wastewater service (for the benefit of Group 2 households only).

4.3.4 Social data

This general information, relating to the characterisation of the primitives of the decision problem faced by the user when he wishes to measure the socio-economic performance of a pricing policy, is then completed by the data of 3 social variables that are:

- the threshold value for the **C**onventional **A**ffordability **R**atio (CAR) above which the Household is detected as facing an affordability problem as defined by the CAR approach;
- the threshold value for the **P**otential **A**ffordability **R**atio (PAR) above which the Household is detected as facing an affordability problem as defined by the PAR approach;
- the (monetary) poverty line defined as the minimum standard of living value below which the household is considered to be a poor household.

The household's standard of living is precisely defined as the ratio between its disposable income (after tax) and an equivalent number of adults measured according to the OECD equivalence scale, with the first adult assigned a coefficient of 1, other adults a coefficient of 0.5 and children under 14 a coefficient of 0.3. The value of this indicator (calculated for each household in the Population module) is interpreted as measuring a level of disposable income for each member of the family. On the basis of the threshold values entered by the user, the tool identifies the situation of each household in the customer file, then breaks down the population into 2 broad categories: the sub-population of households in a situation of poverty and those who are not. This breakdown is used to present the data and calculate several indicators for each of these sub-groups of the population (such as, for example, the percentage of households facing a problem of affordability of the EP / EPA service).

The CAR value corresponds simply to the weight of the water bill (or of the water and wastewater bill if the household is connected to the collective sanitation system) in household’s disposable income while the PAR index corresponds to the value taken by this ratio but for a bill amount calculated for the basic consumption of Household *i*. The values of these budget coefficients are calculated for each household in the Population module, for the IBT and the TBSE tarifs, then compared with the threshold values entered by the user to identify households facing an affordability issue and, where appropriate, to compute the extent of this unaffordability issue. See section 7 for a detailed presentation of this data processing (with the related calculation and breakdown of affordability indicators for the general EP / EPA service).

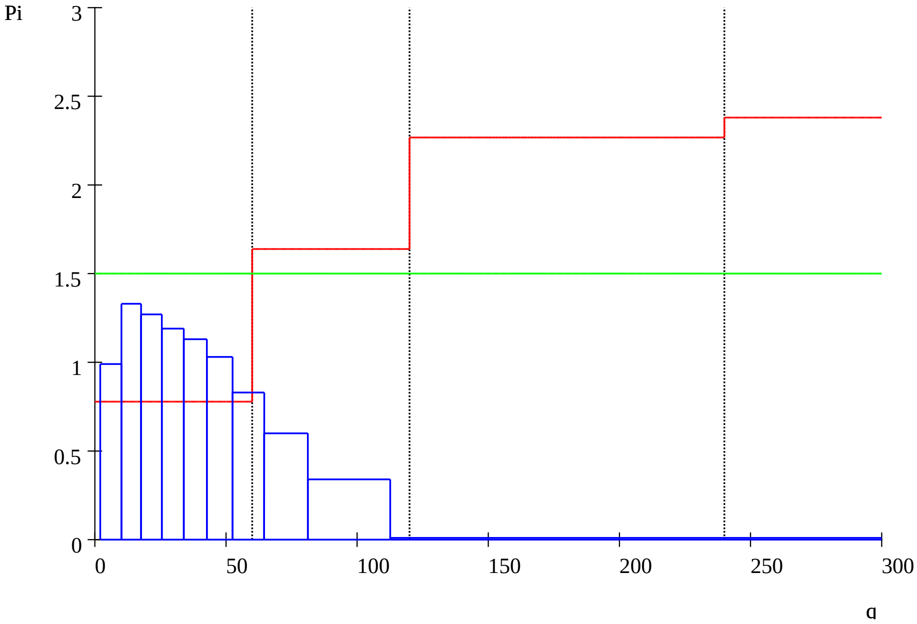


Figure 11 : Subsidy and taxation (rates of) on consumption – EP Tarif 2018 Saint Paul 974 (assuming $c = 1.50€$)

Note for the reader: in blue, the histogram of the distribution of charged water consumption for the Commune of Saint Paul - year 2018.

		T_1	T_2	T_3	T_4	S_q	T_q	$S_q + T_q$	C_0	Δ_DAI
D1	10.1	-7.28				-7.28	0	-7.28	-30	-37.28
D2	17.6	-12.72				-12.72	0	-12.72	-30	-42.72
D3	25.5	-18.41				-18.41	0	-18.41	-30	-48.41
D4	33.9	-24.46				-24.46	0	-24.46	-30	-54.46
D5	42.7	-30.86				-30.86	0	-30.86	-30	-60.86
D6	52.5	-37.88				-37.88	0	-37.88	-30	-67.88
D7	64.5	-43.32	0.63			-43.32	0.63	-42.69	-30	-72.69
D8	81.2	-43.32	2.95			-43.32	2.95	-40.37	-30	-70.37
D9	112.6	-43.32	7.32			-43.32	7.32	-36.00	-30	-66.00
C95	623.0	-43.32	8.34	92.16	337.00	-43.32	437.50	394.18	-30	364.18
C99	1746.8	-43.32	8.34	92.16	1325.96	-43.32	1426.46	1383.14	-30	1353.14
C_type	30.0	-21.66				-21.66	0	-21.66	-30	-51.66
Mean	60.0	-43.32				-43.32	0	-43.32	-30	-51.66

Table 3 : Net subsidies and "taxes" on consumption and access per subscriber, by consumption decile and percentile – EP Tariff 2018 Saint Paul ($c = 1.50€$).

V – DEMAND MODULE

The Demand module calculates and breaks down household water consumption using an econometric model of domestic water demand, in this case the one estimated for Reunion Island, based on the household data in the Population module, by Binet, Carlevaro and Paul [2014] (referred as BCP in the following). These calculations (related to the level and the breakdown of water consumptions) are performed for each household listed in the Population data file, given the water pricing policy which is evaluated/tested by the user.

5.1 Water Demand Econometrics

Using statistical data on household water consumptions, socio-economic characteristics of the household, features of the occupied habitat, tariff parameters …, econometric methods allow estimating Household's water demand functions, that is the causal relationship that links household water consumption to its main determinants (explanatory variables) such as the size and composition of the family, its level of income, the size and type of habitat ... and the characteristics of the water pricing. It should be emphasized that econometric methods most often deal with data that is not derived from controlled experiments (experimental data) and, in so doing, implement specific statistical methods (so-called "observational" studies).

Coupled with the database used for their estimation, this knowledge about the water demand functions of the household provides in turn useful information to carry out some relevant public policy evaluations, whether for diagnostic or simulation purposes. The information in question refers mainly to:

- the measurement of the volumes of water that are necessary to meet the basic needs of the household (and that change across the population),
- the price-responsiveness of water demand

with, for IBTs in particular, the measurement of the impact on household water consumption of (i) a 1 euro increase in the subscription fee, (ii) a 10 centime increase in the price of block 1 (or block 2, block 3, etc.), (iii) a 1% increase in all unit prices (making up the unit price scale) with the "tariff price elasticity", or (iv) an increase in the threshold for block 1 (respectively for block 2, block 3 ...) of 1 cubic metre.

Econometric analysis can also be completed by the implementation of dedicated methods (Generalized Prices of Shin, nested model) to infer perceived prices of the consumers facing potentially complex non-linear pricing schemes, such as IBTs. The latter may undermine the incentive character of pricing, that is its capacity to fix households on sober uses of water, and lead to overconsumptions which it is then possible to measure. See notably Arbuès et al [2003], Worthington & Hoffman [2008] and Montserrat et al. [2015] for reviews of the literature on these aspects of the econometrics of household water demand. The latter deal with the type of data that are used, the specification of the models estimated (choice of functional form, selection of explanatory variables), the statistical methods that are applied to estimate the parameter values (response coefficients) of the model (of household water demand) and the main empirical results obtained in the literature. Reynaud [2015] provides estimates of household water demand functions for each one of the EU-28 countries using aggregate data.

A functional form that has been validated by a large number of empirical studies is the linear expenditure system for which the (conditional) water demand function of household *i* in tranche *j* is of the form:

$$q_{ij}^{d} = \underline{q}_i + \alpha \times \frac{R_i - F + D_j}{\pi_j} - \alpha \times \frac{p_2 \underline{q}_{i2}}{\pi_j} \qquad (5.1)$$

with[11]:

- F the fixed part (subscription amount),
- D_j the value of Nordin D in block *j* ("virtual rebate"),
- π_j the unit price (per cubic metre) of block *j* ("marginal price"),

(see section 3.2) and:

- R_i is the income of household *i*,
- $\underline{q}_i$ its basic consumption of tap water,
- $\underline{q}_{i2}$ its basic consumption for the "other goods" (composite good),
- p_2 the price of the "other goods",
- $\alpha \in [0,1]$ a preference parameter.

The term "linear" comes from the fact that it is associated with the demand function (5.1) an expenditure function for block *j*:

$$T_{ij} = T\left(q_{ij}^{d}\right) = (1-\alpha)\left(F - D_j + \pi_j \underline{q}_i\right) + \alpha \times \left(R_i - p_2 \underline{q}_{i2}\right) \qquad (5.2)$$

which has the particularity of varying linearly with income (it is precisely this relationship (5.2) which is estimated, using appropriate statistical methods, with estimated values for the coefficients of the water demand function that can be found by identification).

As tap water is little consumed on its own, a demand function such as (5.1) should be seen more as giving the volume of water that will be used by the household, taking into account in particular its level of income and the characteristics of the tariff, as a quantity of input in a domestic production function, with the production of an output that can be defined, in the broadest sense, as being "Family Well-Being" (not necessarily measurable).

The BCP functional form implemented in the micro-simulation model postulates a relationship similar to the equation (5.1) with variables (consumption, basic consumption, variable part, etc.) expressed not in levels but in logarithms.

[11] These 3 variables relating to the "EP" or "EPA" tariff system faced by household *i*, depending on whether or not it is connected to the public sewerage network, must be understood to include all taxes and charges.

5.2 The function implemented in the DST

5.2.1 Captive component vs. "economic" component

The demand function implemented in the DST corresponds to the econometric model (in fact, to its deterministic part) estimated by BCP [2014] on a sample of households residing on Reunion Island. This model breaks down the logarithm of household water consumption into the sum of 2 components:

$$\ln q = \ln q_0 + \ln f\left(R, F, \pi_1, \ldots, \pi_p, k_1, \ldots, k_{p-1}, \kappa\right) + \varepsilon \quad (5.3)$$

with:

- ε the error term
- $\ln q_0$ the captive part (of the logarithm) of demand
- $\ln f(\cdot)$ the variable part (also sometimes called the "economic" part) of this same demand function (in logarithm).

By analogy with the linear system of expenditure, the captive portion gives the proportion (of the logarithm) of consumption that does not respond to variations in household income and/or changes in tariff parameters. It therefore takes the form of an incompressible consumption (or perceived as such by the household), linked to domestic uses of water under normal conditions of household activity (production of domestic goods and services) and for which consumption habits (more or less efficient use of water-consuming equipment) also play a role. Following Gaudin et al [2001], this captive part is in turn modelled as a function of the size of the family, its composition and its equipment geared towards outdoor uses of the resource by means of the relationship:

$$\ln q_0 = -2.56 + 0.48 \cdot \ln N + 0.44 \cdot \text{SNWA} + 0.12 \cdot \text{Pool} + 0.37 \cdot \text{Garden} \times \text{Weather} \quad (5.4)$$

with N the family size, SNWA the share of non-working adults with respect to the total number of adults within the household, Pool a dummy variable which takes the value 1 if the household has a swimming pool (0 otherwise), Garden a dummy variable which takes the value 1 if the household has a garden (0 otherwise), Weather the percentage of days without rain over the billing period (rainfall frequency). The coefficients appearing in front of each of these variables (and whose values can be modified by the user) have an economic significance, which is explained later in the text. The variable part is modelled as follows:

$$\ln f = -0.31 \ln\left(\bar{\pi}_j^{1-\kappa} \times \pi_j^{\kappa}\right) + 0.25 \ln\left[\left(R - F + D_j\right)^{\kappa} \times \left(R - F\right)^{1-\kappa}\right] \quad (5.5)$$

with π_j the unit price of the block *j* in which the household's consumption places it (marginal price), D_j the value of the associated Nordin D ("virtual reimbursement"), $\bar{\pi}_j$ the average price calculated excluding subscription fees for this same consumption $q \in I_j$:

$$\bar{\pi}_j = \frac{T(q) - F}{q} = \pi_j - \frac{D_j}{q} \quad (5.6)$$

and $\kappa \in [0,1]$ a tariff perception parameter whose value estimated by BCP [2014], $\hat{\kappa}$, is not significantly different from 0. This modelling generalises the specification in terms of perceived price developed by Shin [1985] by also involving the parameter κ on the measure of "virtual" income of block *j*, with a consumer who perceives properly the tariff when the parameter κ is close to 1 (Nordin specification) and who reasons in terms of average price when the parameter κ is close to 0 (Taylor specification). As with the captive part of consumption, the values of the coefficients appearing in front of these different price and income variables can be modified by the user (the value of the perception parameter κ must be entered by the user in order to run a simulation) and have an economic significance that is explained later in the text.

5.2.2 Estimating basic needs

5.2.2.1 The captive component – properties

As emphasized above, the captive component of the water demand relates to the part of the water consumption which is not affected by changes in prices or income. The latter varies nevertheless with family size, water-using devices, habitat characteristics ... according to the formula given by (5.4) with estimated values for the response coefficients (of the various explanatory variables (regressors)) that have a socio-economic meaning.

- The coefficient $a_1 = 0.48$ attached to the logarithm of family size measures the elasticity of household water consumption in relation to family size.

The value obtained indicates that, all other things being equal, a 1% increase in family size is accompanied by a 0.48% increase in household water consumption. This highlights a positive but non-linear effect of family size on consumption (an increase from $N_i = 1$ to $N_i = 2$ corresponds to an increase of 100%, and therefore a variation in consumption of 48%; an increase from $N_i = 2$ to $N_i = 3$ to an increase of 50%, and therefore an increase in consumption of 24%, etc.), which is due to collective element and existence of economies of scale in domestic water use.

Note A more satisfactory way of interpreting the value of this coefficient is to think in individual terms with a comparison, all other things being equal, of the water consumption of households of different sizes. In this context, the model states that a 2-person household consumes, on average, $0.48 = 48\%$ more water than a single-person household with the same socio-economic characteristics. Similarly, the water consumption of a 3-person household is, on average, 24% higher than that of a 2-person household with the same characteristics in terms of the other determinants of demand and so on.

- The SNWA variable refers to a notion of presence at home.

The positive value (0.44) obtained by BCP [2014] for the estimation of the coefficient of this variable establishes that, all other things being equal (and therefore for a fixed family size), a household's water consumption is all the greater the lower the employment rate within the family, with impacts that are then quite strong. Compared with a couple of adults who are both in employment and for whom the SNWA variable takes the value 0, the value 0.44 means that the household's water consumption is (to a first approximation) 44.0% higher when the two adults are not in employment, and 22.0% higher when one of the two adults is not in employment (a difference in logarithms approximates a growth rate).

- The estimated value (0.12) of the coefficient associated with the Pool dummy variable measures the impact on household water consumption of owning a swimming pool.

All other things being equal (same family size, same composition, same level of income, etc.), households with a swimming pool consume 12% more water on average than households without one.

- The product of the variables Garden (which is a dummy variable) and Weather (the frequency of rainfall over the billing period) shows that climate has an impact on domestic water consumption (in the Population module) but only for households with a garden (this variable takes the value 0 when the indicator variable Garden takes the value 0, i.e. when households do not have a garden).

In this context, the coefficient 0.37 indicates that, assuming that it does not rain over the entire billing period, a household with a garden consumes on average, all other things being equal, 37% more than a household without a garden[12]. As the frequency of rainfall is in fact non-zero (58% on average for the households in the initial sample), the presence of this factor makes it possible to calculate for each of the household in the Population module a consumption that can be attributed to having a garden and which appears, in fact, to be a fairly high consumption factor (+21.4% on average)[13] .

Last, coupled with the realisations of the error term ε, the constant -2.56 measures the effects of the other variables (omitted variables) on the logarithm of consumption.

5.2.2.2 Distribution of basic needs - high estimate

Taken together, these variables lead to an estimate of captive consumption, which varies from household to household and whose distribution is shown in the histogram on Figure 12, page 58. The values obtained are therefore quite significant with:

(i) a "fixed" water consumption that sets, on average, to 463.5 litres per day per household, (around 169 m³ on the basis of one year's consumption),

(ii) consumption quartiles estimated to 300, 450 and 620 litres per day and per household, (that is 109, 164 and 226 m³ per year per household)[14].

---

[12] Since the interpretation is made ceteris paribus, this figure differs from the one that would be obtained by comparing the water consumption of households equipped with a garden versus the water consumption of households not equipped with a garden (with socio-economic composition that differs a priori).

[13] Furthermore, as in Schleich and Hillenbrand [2009], the insignificance of other climate variables (reported by BCP [2013]), such as the amount of rainfall or the level of evapotranspiration, suggests that households (in the Population file) base their garden watering decisions not on the level of rainfall but on the simple fact that it has rained or not, with a consequent misuse of water in household management of this amenity.

[14] Literally, 25% of households have captive consumption of less than 109 cubic metres per year, 50% of less than 164 cubic metres per year and 75% of households of less than 226 cubic metres per year. In this way, are identified 4 consumption intervals (less than 109 cubic metres per year, between 109 and 164 cubic metres per year, between 164 and 226 cubic metres per year, more than 226 cubic metres per year), each comprising a quarter of the population of domestic subscribers (households)..

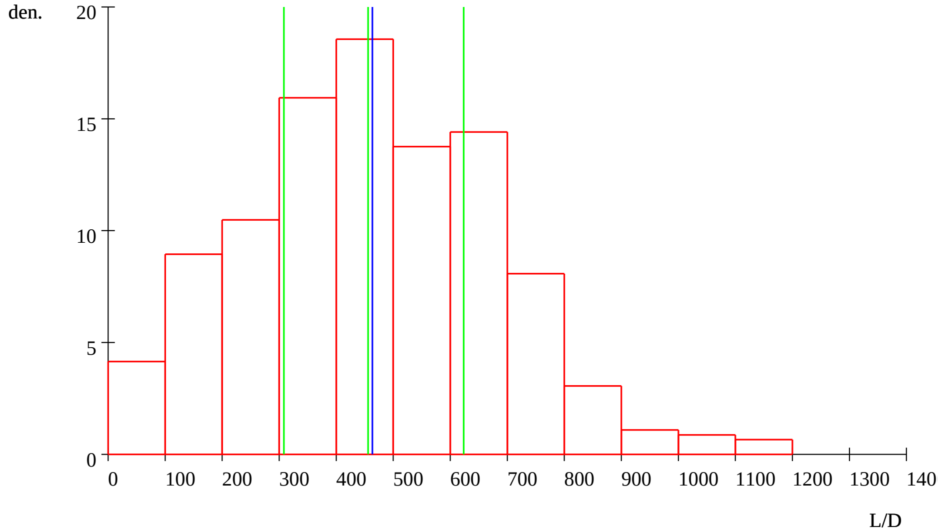


Figure 12 : Distribution of captive consumption - BCP [2014]

Following Gaudin et al [2001], these captive consumptions constitute in turn a high estimate of the volumes of water necessary to cover the basic needs of households and the knowledge (in fact, the estimate) of this distribution (as provided by Figure 12) makes it possible to assess the correct calibration of the tariff with, in particular, the measurement of inclusion and exclusion errors in volume (see paragraph 7.3) and in value (see paragraph 8.1.3) which are implemented by an IBT. As highlighted earlier in the text, the design of the tariff scale involves subsidies on the first consumption blocks and "taxes" on the higher blocks, which are collected in order to balance the funding of the cross-subsidy system. Thus, for households in the (initial data file of the) Population module, an initial block of 40 m³ per quarter covers the basic needs (estimated at a high level by the captive component) of roughly half the sample, and this rate rises to 75% for a threshold of around 60 m³ per quarter. Examination of the characteristics of the remaining 25% of households that would potentially be subject to non-social pricing of the service for part of their basic consumption shows that (*i)* these households are heavy consumers (957 l/d per day per subscriber on average) with, at the same time, (*ii*) much lower incomes (around 30%), (*iii*) more adults who are not in employment (2.46 compared with 0.91 for the sample average), (*iv*) a larger family size (1 person more on average) and (*v*) almost systematic use of a garden (97%). These factors therefore show that, by setting the thresholds for the first blocks, the water manager does indeed face a risk of exclusion of large poor families for household in the Population module.

5.2.2.3 Retreatment of the captive part

This potential risk of exclusion of large poor families must, however, be put into perspective by the fact that captive consumption q_0 includes as determinants the Garden variable and also the Pool variable, which may be considered to fall outside the scope of basic needs. For this reason, the tool:

(1) provides an estimate of the basic consumption of households in the Population module, denoted:

$\underline{q}_1, \underline{q}_2, \ldots, \underline{q}_n$

based solely on the variables N (family size) and SNWA (rate of non-employment within the family) using the relationship:

$$\ln \underline{q} = -2.56 + 0.48 \cdot \ln N + 0.44 \cdot \text{SNWA} \qquad (5.7)$$

(for the default values of the response coefficients),

(2) offers the user the possibility of adding to the list of determinants of basic consumption one and/or other of these two variables, Garden and Pool, which also enter into the determination of captive consumption, so as to obtain estimates of basic consumption that are more in line with the qualification of the various uses of water (basic, comfort, luxury) the user may have (the discussion may focus in particular on the Garden, part of which may be devoted to food production with the management of a vegetable garden).

On this basis, the model computes for each household in the Population module :

(3) the volume of non-basic water consumption $q - \underline{q}$

(4) the volume of the variable part of water consumption $q - q_0$

by subtracting from the volume of consumption returned by the water demand function $q_i = q_i^d(\cdot)$ (predicted value) the level of its basic consumption $\underline{q}$ and that of its captive part q_0.

5.2.3 Tariff perception

The second component of the demand function (*see* eq. (5.3) page 55) gives the part (of the logarithm) of the household's consumption that responds to variations in its income, on the one hand, and to a change in the IBT pricing parameters, on the other. This is the part of demand that can be targeted for water savings in the short term, by means of an appropriate pricing policy[15]. As mentioned above, the parameter κ is a tariff perception parameter whose value must be entered by the user and which, through the polar cases $\kappa = 0$ and $\kappa = 1$, accounts for two main types of behaviour: the Taylor consumer who reasons in terms of average price (Taylor [1975]) and the Nordin consumer who reasons in terms of marginal price (Nordin [1976]).

5.2.3.1 The polar case $\kappa = 1$

When $\kappa = 1$, tariff perception is perfect with (conditional) demand of tranche *j* for household *i* that writes as[16]:

[15] The demand function estimated by BCP [2014], by linking domestic water consumption to a given level of the household's water-consuming equipment, constitutes a short-term demand (by definition). At the same time, and similarly to the capital demand of a firm, variations in the price of water are likely to lead to a change in the quantity and/or in the quality of the household's equipment, which then has an impact on its water consumption. The demand function that takes account of this induced change in equipment is referred to as long-term demand. Most empirical studies that estimate econometric model of household water demand estimate short-term demand. There are, however, several contributions that estimate long-term demand (which is more demanding in terms of data with the use of panel data). See notably Nauges and Thomas [2003] and Almendarez-Hernández et al [2016].

[16] For ease of reading, the discussion is conducted using the estimates obtained by BCP [2014]. It should be kept in mind that the user has the option of modifying the values of these response coefficients in the Demand module.

$$\ln q_i = \ln q_{0i} - 0.31 \ln \pi_j + 0.25 \ln \left(R - F + D_j \right) \tag{5.8}$$

with, as a reminder, q_{0i} the captive consumption of household *i*, R_i the level of its quarterly income, π_j the unit price (per cubic metre) of block *j* and D_j the value of Nordin's D ("virtual reimbursement") of block *j*. This demand function can be compared with that which emerges with a two-part tariff with parameters (F, π) for which[17]:

$$\ln q_i = \ln q_{0i} - 0.31 \ln \pi + 0.25 \ln \left(R_i - F \right) \tag{5.9}$$

Under these conditions, it appears that, for consumption q_i located in block *j*, all goes as if (*i*) the price were given by the marginal price π_j, that is the price of the consumption block in which the household is located, and (*ii)* the household income were not given by R_i but by $R_i - F + D_j$, that is the household income net of the amount of the subscription fee but increased by the value of the Nordin D of block *j* ("virtual income").

These equivalences (in the measurement of price and income), linked to the accounting breakdown of the water bill given by the equation (4.9) on page 36, provide a reading grid from which it is relatively simple to analyse the (local) impacts of the various pricing parameters on household water consumption. For the main:

- the price of the consumption block in which the household locates (marginal price) plays the role of a traditional price which then has a negative impact, in the usual way, on the household's water consumption;
- the inframarginal characteristics of the tariff (prices and thresholds of the consumption blocks that precede the consumption block in which the household locates, if any), play a role through an income effect, via the modification of the Nordin D and that of the "virtual" income $R - F + D_j$ thus generated;
- the supramarginal characteristics of the tariff (prices and thresholds for consumption blocks past the consumption block in which the household locates, if any) play no role (locally).

See Appendix A for the identification of these effects (and the calculation of the impacts) linked to changes in the tariff parameters of an IBT4 for water consumption by households in blocks 2 and 3. This specification accounts for the behaviour of a consumer who has fully understood the properties of a progressive pricing system and, as a result, implements a fully optimal management of his domestic water use. The latter is characterised by an equalisation, at the optimum, between (*i*) the household marginal cost of water consumption as given by the marginal price scale $\partial T / \partial q = \pi(q)$, and (*ii*) the household marginal gain in water consumption / the household's willingness to pay to increase its water consumption by one unit (which decreases with the volume of water consumed q_i and increases with the consumer's income R_i). This type of consumer is referred to in the literature as the Nordin consumer.

[17] Note that a two-part tariff of parameters (F, π) can be analysed as an IBT2 of parameters (F, π_1, π_2, k_1) in which $\pi_1 = \pi$ and $k_1 \to +\infty$. Under this last condition, it is understood that household water consumptions will all be in block 1 with a (conditional) block 1 demand which is then given by (4.9).

5.2.3.2 The polar case $\kappa = 0$

In this second polar case, the demand model is of the form:

$$\ln q_i = \ln q_{0i} - 0.31 \ln \bar{\pi}_j + 0.25 \ln \left(R_i - F \right) \tag{5.10}$$

where (as a reminder) $\bar{\pi}_j$ is the average price (calculated excluding subscription fee) for consumption q in block *j*:

$$\bar{\pi}_j = \frac{T(q) - F}{q} = \pi_j - \frac{D_j}{q} \tag{5.11}$$

Comparing this demand model (consisting of equations (5.10) and (5.11)) with the one that emerges with two-part tariff (*cf.* equation (5.9) on page 60), it is seen that:

- the measure of income is now made up of income net of the subscription fee, similar to a two-part tariff,

and:

- the measure of price is now made up of the average variable cost of consumption (for consumption q in block *j*).

This specification corresponds to the behaviour of a consumer who, faced with a progressive pricing system that may present a certain degree of complexity, summarises the properties of the unit price scale by computing an average price, then optimises his water consumption by equalising his marginal willingness to pay for water (see above) at this average (or perceived) price level $\bar{\pi} \leq \pi_j$.

This way of dealing with the unit price scale is a heuristic known as "ironing" (*cf.* Liebman & Zeckhauser [2004][18]) and generates (except in the specific case of consumption in block 1; see below) sub-optimal management of consumption by the household, with:

- over-consumption linked to an underestimation of the marginal cost of consumption

and

- mismanagement costs linked to the fact that the household reasons in terms of average price and not in terms of marginal price (cognitive bias).

This type of consumer is referred to in the literature as a Taylor consumer.

Note It should be noted that the consumption decision is always based, when $\kappa = 0$, on a logic of optimisation, but with a poor measure of the marginal cost of consumption. For the main, the household substitutes the IBT (and the unit price scale $\pi = \pi(q)$ he faces) with a flat tariff corresponding, at the stationary equilibrium, to the long-run value of the average cost of consumption (this flat tariff is therefore an inaccurately perceived schedule that is called a "schmedule" by Lienman & Zeckhauser, op. cit.). *See* Paul [2023] for a presentation and study of this dynamic model.

[18] As stated by the authors, "Ironing arises when an individual facing a multipart schedule perceives and responds to the average price at the point where he consumes".

5.2.3.3 The BCP specification

Based on these two polar cases $\kappa=0$ and $\kappa=1$, the BCP specification [2014] is initially a nested (statistical) model which simply consists of combining the Nordin and Taylor models by means of a nesting parameter κ:

$$\ln q_i = (1-\kappa)\ln q_i^{\kappa=0} + \kappa \ln q_i^{\kappa=1} \tag{5.12}$$

with:

$$\ln q_i^{\kappa=0} = \ln q_{i0} - 0.31 \ln \bar{\pi}_j + 0.25 \ln (R_i - F)$$

$$\ln q_i^{\kappa=1} = \ln q_{i0} - 0.31 \ln \pi_j + 0.25 \ln (R_i - F + D_j)$$

to obtain in the end:

$$\ln q_i = \ldots = \ln q_{i0} - 0.31 \times \ln\left(\bar{\pi}_j^{1-\kappa}\pi_j^{\kappa}\right) + 0.25 \times \ln\left((R_i - F)^{1-\kappa}(R_i - F + D_j)^{\kappa}\right) \tag{5.13}$$

On the basis of this, the closer the parameter κ is to 0, the closer consumption behaviour will be to the Taylor consumer, and the closer the parameter κ is to 1, the closer consumption behaviour will be to the Nordin consumer.

This (ad hoc) specification makes it possible to generalise the modelling proposed by Shin (op. cit.), which consists of including a weighted geometric mean of the marginal price and the average price, $\pi^* = \bar{\pi}^{1-\kappa}\pi^{\kappa}$, as a measure of price, which is then called the "perceived price". The main limitation of Shin's modelling is that it does not allow to recover the Nordin specification when $\kappa=1$, no correction being made at the same time to the income measure. The BCP specification [2014] avoids this drawback of the Shin's approach.

Additional Information The user having the possibility, as with the determinants of captive consumption, of modifying the response coefficients attached respectively to the logarithm of the perceived price and the logarithm of the perceived income (that comes with), the following (4) properties will complete this presentation.

P1 Since the average price is calculated excluding subscription fees, the BCP specification does not generate overconsumption for households facing an IBT with consumption q_i located in block 1.

Details In this case, the average price is equal to the marginal price:

$$\bar{\pi}_1 = \pi_1 - \frac{D_1}{q} = \pi_1 - \frac{0}{q} = \pi_1 \tag{5.14}$$

with a perceived price equal to $\bar{\pi}_1^{1-\kappa}\pi_1^{\kappa} = \pi_1^{1-\kappa}\pi_1^{\kappa} = \pi_1$ (marginal price of tranche 1) and a demand function given (locally) by:

$$\begin{aligned}\ln q_i &= \ln q_{0i} - 0.31 \times \ln\left(\bar{\pi}_1^{1-\kappa}\pi_1^{\kappa}\right) + 0.25 \times \ln\left((R_i - F)^{1-\kappa}(R_i - F + D_1)^{\kappa}\right) \\ &= \ln q_{0i} - 0.31 \times \ln\left(\pi_1^{1-\kappa}\pi_1^{\kappa}\right) + 0.25 \times \ln\left((R_i - F)^{1-\kappa}(R_i - F + 0)^{\kappa}\right) \\ &= \ln q_{0i} - 0.31 \times \ln \pi_1 + 0.25 \times \ln (R_i - F)\end{aligned} \tag{5.15}$$

P2 Property P1 also applies to two-part tariffs (F,π) for which:

$$\bar{\pi} = \frac{T(q)-F}{q} = \frac{F+\pi q-F}{q} = \pi \tag{5.16}$$

with a demand equation given by:

$$\ln q_i = \ln q_{0i} - 0.31 \times \ln \pi + 0.25 \times \ln\left(R_i - F\right) \tag{5.17}$$

P3 For a given value of the perception coefficient κ, the coefficient (here -0.31) of the logarithm of the perceived price $\bar{\pi}_j^{1-\kappa}\pi_j^{\kappa}$ in the conditional demand for tranche *j*:

$$\ln q_i = \ln q_{i0} - 0.31 \times \ln\left(\bar{\pi}_j^{1-\kappa}\pi_j^{\kappa}\right) + 0.25 \times \ln\left(\left(R_i - F\right)^{1-\kappa}\left(R_i - F + D_j\right)^{\kappa}\right) \tag{5.18}$$

measures initially (i) the elasticity of water consumption in relation to the price of a cubic metre when faced with a two-part tariff (F,π):

$$\eta_{q_i^d,\pi} = \frac{\partial \ln q_i^d}{\partial \pi} = \frac{\partial q_i^d}{\partial \pi}\frac{\pi}{q_i^d} = \ldots = -0.31 \tag{5.19}$$

i.e. a 1% increase in the price (per cubic metre) π reduces household water consumption by 0.31%, and also (ii) the elasticity of water consumption of a household located in block 1 when faced with an increase in the block 1 price:

$$\left.\frac{\partial \ln q_i^d}{\partial \ln \pi_1}\right|_{q_i^d \in I_1} = \left.\frac{\partial q_i^d}{\partial \pi_1}\right|_{q_i^d \in I_1}\frac{\pi_1}{q_i^d} = -0.31 \tag{5.20}$$

i.e. a 1% increase in the price of block 1 reduces water consumption by household located in block 1 by 0.31%. These demand functions then have the particularity of being inelastic with respect to the price π / the (marginal) price π_1 of tranche 1 (in accordance with the empirical results highlighted in the literature)[19].

P4 Back to the conditional demand equation (5.3) & (5.5), the coefficient (here 0.25) of the logarithm of the perceived "virtual" income $\left(R_i - F\right)^{1-\kappa}\left(R_i - F + D_j\right)^{\kappa}$ measures nearly the elasticity of water consumption in relation to income when the household is facing a two-part tariff (F,π) with[20]:

$$\eta_{q_i^d,R} \equiv \frac{\partial \ln q_i^d}{\partial \ln R} = \frac{\partial q_i^d}{\partial R}\frac{R}{q_i^d} = \ldots = 0.25 \times \frac{R_i}{R_i - F} \approx 0.25 \text{ pour } R >> F \tag{5.21}$$

[19] Meta-analysis of price elasticities are provided by Dalhuisen et al [2003], Sebri [2012] and Marzano et al [2018]. The first 2 references analyse as well income elasticities of residential water demand. See also Grafton et al. [2009].
[20] Note that, for a two-part tariff, the perceived price is equal to π (as explained above) with, at the same time; a Nordin's D that is equal to 0.

i.e. the water consumption of a household facing a two-part tariff with parameters (F,π) increases (approximately) by 0.25% when its income increases by 1%[21]. Water consumption appears to be a 'superior' good, i.e. water consumption increases with household income, but not a luxury good, i.e. the share of income devoted to paying the water bill decreases with household income. These properties are consistent with the empirical results highlighted in the literature.

It should also be noted that, since demand functions are (from a theoretical point of view) homogeneous of degree 0 in price and income (a property known as the "No monetary illusion"), the value of the coefficient used to measure income elasticity (here 0.25) should normally be equal, in absolute terms, to the value of the coefficient used to measure price elasticity (here -0.31). It is up to the user to decide whether or not to align these 2 values.

5.3 Other issues

(1) The demand function (5.12) given on page 62 is derived from a (theoretical) model in which the household programmes at each date *t* a consumption flow for the period $\left[t,t+1\right]$ (so-called planned consumption) on the basis of the information provided by its last bill at date *t* and which relates to the previous consumption period, that is the time interval $\left[t-1,t\right]$. This dynamic model is presented and analysed in Paul [2023]. The main point is that the time series of consumption thus generated are locally stable in the neighbourhood of the stationary equilibrium with a related dynamic of damped oscillations.

The tool then manages this aspect in the following way.

- For a given value of κ (the value entered by the user), the model first calculates the household's consumption and the amount of the water bill on the assumption that the tariff is properly perceived, what provides an initial point.

Then:

- In a second step, one calculates an average price (excluding subscription fee) and recalculates the household's water consumption by including this average price value in the demand function (5.3) & (5.5).

This operation results in an increase in consumption (for all households in block 2 and above), an increase in the bill and a new average price, which is again fed into the demand function (5.3) & (5.5) and so on. The tool then displays the value of consumption at a date *t* in this process (which essentially amounts to considering that the consumer has suddenly started to think incorrectly) and which corresponds to the value entered by the user (in the Demand module). The user can examine the results over a time horizon ranging from 1 to 8 periods, i.e. from 1 quarter to 2 years. In practice, the convergence is rather rapid with an assessment that can be made at $T=4$ quarters (that is, after one year of consumption).

---

[21] As the billing period is quarterly, the same applies to water consumption and also to household income, which appears in the demand function (4.18). Numerically, the value of this quarterly income R_i is then effectively greater than the amount of the subscription F with a ratio $R_i / R_i - F$ that is very close to 1 (for example, for a household income of 2000 euros net per month, i.e. 6000 euros per quarter and a subscription at 30 euros per quarter, the ratio $R / R - F$ takes the value 1.005).

(2) Calculating consumption when the tariff is perfectly perceived ($\kappa = 1$) also requires an algorithm to be implemented, since the equation (5.8) is not the equation of the "complete" demand function which is here of the form:

$$\ln q_i^d = \begin{cases} \ln q_{0i} - 0.31 \ln \pi_1 + 0.25 \ln(R_i - F) & if \quad R_i \le \bar{R}_{i1} \equiv F + \left(\frac{k_1 \pi_1^{0.31}}{q_{0i}}\right)^{\frac{1}{0.25}} \\ \ln k_1 & if \quad \bar{R}_{i1} \le R_i \le \underline{R}_{i2} \equiv F - D_2 + \left(\frac{k_1 \pi_2^{0.31}}{q_{0i}}\right)^{\frac{1}{0.25}} \\ \ln q_{0i} - 0.31 \ln \pi_2 + 0.25 \ln(R_i - F + D_2) & if \quad \underline{R}_{i2} \le R_i \le \bar{R}_{i2} = F - D_2 + \left(\frac{k_2 \pi_2^{0.31}}{q_{0i}}\right)^{\frac{1}{0.25}} \\ \ln k_2 & if \quad \bar{R}_{i2} \le R_i \le \underline{R}_{i3} \equiv F - D_3 + \left(\frac{k_2 \pi_3^{0.31}}{q_{0i}}\right)^{\frac{1}{0.25}} \\ \ln q_{0i} - 0.31 \ln \pi_3 + 0.25 \ln(R_i - F + D_3) & if \quad \underline{R}_{i3} \le R_i \le \bar{R}_{i3} \equiv F - D_3 + \left(\frac{k_3 \pi_3^{0.31}}{q_{0i}}\right)^{\frac{1}{0.25}} \\ \vdots \end{cases}$$

(with the numerical values obtained by BCP [2014]) but that of the conditionnal demand function of block *j*, for *j* varying from 1 to *p*. The procedure (which applies for each household *i*) is then as follows.

Step 1 First, the level of consumption of household *i* is calculated assuming that it is in block 1, i.e. using the block 1 conditional demand function:

$$\ln q_i = \ln q_{i0} - 0.31 \ln \pi_1 + 0.25 \ln(R_i - F) \tag{5.22}$$

and one examines whether this quantity q_{i1}^* is less than k_1. If this is the case, one sets $q_i^d = q_{i1}^*$ and stops (because this level of consumption corresponds effectively to the optimal consumption of agent *i*). Otherwise, the solution q_{i1}^* is not admissible and one moves on to the next stage of the algorithm.

Step 2 The level of consumption of household *i* is calculated assuming that it is in block 2, i.e. using the block 2 conditional demand function:

$$\ln q_i = \ln q_{i0} - 0.31 \ln \pi_2 + 0.25 \ln(R_i - F + D_2) \tag{5.23}$$

and one examines whether this quantity q_{i2}^* is less than k_1. If so, one stops and sets $q_i^d = k_1$ (this corner solution describes a situation in which the household, given the price increase after the tariff threshold k_1, has no interest in increasing its consumption and going beyond k_1). If not, one moves on to the next stage of the algorithm.

Step 3 One checks whether the consumption q_{i2}^* is less than k_2 and constitutes an acceptable solution (at this stage). If so, one sets $q_i^d = q_{i2}^*$ and stops (because this level of consumption corresponds effectively to the optimal consumption of agent *i*). If not, the solution q_{i2}^* is not admissible and one moves on to the next stage of the algorithm.

Step 4 The level of consumption of household *i* is calculated assuming that it is in block 3, i.e. using the block 3 conditional demand function:

$$\ln q_i = \ln q_{i0} - 0.31 \ln \pi_3 + 0.25 \ln \left(R_i - F + D_3 \right) \quad (5.24)$$

and one examines whether this quantity q_{i3}^* is less than k_2. If so, one stops and sets $q_i^d = k_2$ (corner solution). If not, one moves on to the next stage of the algorithm, which consists of checking whether $q_{i3}^* < k_3$ (in which case this quantity q_{i3}^* is admissible) and so on.

To conclude on this point, it should be noted that (*i*) the demand function resulting from a problem of maximising a quasi-concave objective function ("utility function") in a convex set, the solution is known to exist and to be unique (this is the reason why one is allowed to conclude in the various cases listed above). On the other hand, (*ii*) while the calculation of the complete demand function is analytically possible here, with in particular the determination of explicit expressions for the threshold incomes for which the solutions q_{i1}^*, k_1, q_{i2}^*, k_2, q_{i3}^* ... apply, the same may not be true for other specifications of the water demand function with, in this case, a resolution that has to be done numerically (taking into account the values for the tariff parameters entered by the user). The implementation of this algorithm makes it possible to dispense with this numerical calculation of the complete demand for determining household consumption. This point also applies to the calculation of long-term demand (stationary equilibrium) when the perception parameter κ is not equal to 1.

(3) It should be borne in mind that the estimate of the demand function resulting from the use of econometric methods is by no means universal and must be understood in the local sense of the term. Notably, there is no reason why the water demand functions of households living in the Spanish municipality of Figueres, one of the demo sites for the InnWater project, should be the same as those of households living on Reunion Island. In particular, it is possible (not to say probable) (i) that the explanatory variables (entering into the demand function) are not the same, or even (ii) that the response coefficients (for a given set of regressors) differ.

For the main, a demand function describes a behavioural function and, when it comes to drinking water consumption, it is understood that consumption habits (themselves linked to socio-cultural factors and the relationship of population to water) will play an role in this respect. For this reason, if the tool is to be truly satisfactory and fully operational, the econometric model of local household demand for Reunion Island (estimated by BCP [2014]) must be replaced by an econometric model of local demand for which the user intends to analyse the performance of the local pricing policy. In this sense, the tool is at this stage a prototype that can be used for demonstration and training purposes (academic and professional), in the manner of a Serious Game, to enhance the skills of students (particularly those studying Economics, Management of Public Organisations and Water Sciences) and stakeholders (involved in the decision-making process relating to the setting of tariffs for public drinking water and sanitation services).

(4) In conclusion, it should be noted that the (multiplicative) specification used for the BCP demand model [2014] differs significantly from the linear expenditure model commonly used in the literature (see above). However, the code used to assess the socio-economic performance of the pricing policy was written taking household water consumption and its breakdown (captive vs. variable part, basic vs. non-basic consumption) as inputs. In this way, the adaptations required to enable the tool to evaluate the performance of the pricing policy in the more usual case of a linear expenditure system are slight.

VI –INVOICES MODULE

The Invoices module brings together a range of household water accounts including a Consumption section, an Expenditures section, a Service Costs section, a Subsidies-Taxes section and a Surplus section.

4.1 Consumption Heading For each household in the Population module, information is provided firstly on consumption including:

- the level of its captive consumption q_{0i} ,
- the level of its basic consumption $\underline{q}_i$ (considering the user's reprocessing),
- the captive part of its consumption that does not meet basic needs, in this case water used for garden maintenance and swimming pool maintenance (excluding user reprocessing), $\underline{q}_i - q_{0,i}$,
- the level of its IBT consumptions for (i) the value of the perception parameter $\kappa = \kappa_0$ entered by the user, q_i^{IBT} , and (ii) the reference case $\kappa = 1$ (with a tariff that is properly perceived), $q_i^{\text{IBT-PP}}$ (the acronym PP stands for 'Perfection Parfaite' in French), given the IBT tariff which is evaluated/tested by the user,
- the level of overconsumption $q_i^{\text{IBT-PP}} - q_i^{\text{IBT}}$, linked to a poor perception of the tariff, equal to 0 (no overconsumption) if the Household is a small consumer in block 1 and/or if there is no poor perception of the tariff (with a parameter κ equal to 1),
- the level of its TBSE consumption, q_i^{TBSE} , based on the values entered by the user for the cost parameters (fixed cost *CF*, number of subscribers *n*, variable unit production cost *c*), the VAT rate and the unit excise dutie (General Data tab of the micro-simulation model),
- the level of non-basic consumption for (i) IBT pricing when the latter is poorly perceived, $q_i^{\text{IBT}} - \underline{q}_i$, (ii) IBT pricing when the latter is properly perceived $q_i^{\text{IBT-PP}} - \underline{q}_i$, and (iii) TBSE pricing, $q_i^{\text{TBSE}} - \underline{q}_i$,
- the level of the variable part (economic component) of (i) IBT consumption $q_i^{\text{IBT}} - q_{0,i}$, (ii) IBT-PP consumption $q_i^{\text{IBT-PP}} - q_{0,i}$, and (iii) TBSE consumption $q_i^{\text{TBSE}} - q_{0,i}$,

and the breakdown by block of all of these consumption variables with for instance:

(i) a basic consumption for block 1, $\underline{q}_i^1$;

(ii) a basic consumption for block 2, $\underline{q}_i^2$;

…

(p) a basic consumption for block *p*, $\underline{q}_i^p$;

(and $\underline{q} = \underline{q}_i^1 + \underline{q}_i^2 + \ldots + \underline{q}_i^p$) for each of the consumption blocks set by the user (with some of these variables potentially equal to 0). This dataset also includes block variables :

$\underline{\mathrm{B}}_i$, B_i^0, $\mathrm{B}_i^{\mathrm{IBT}}$, $\mathrm{B}_i^{\mathrm{IBT\text{-}PP}}$

indicating the number of the consumption block in which the basic consumption, captive consumption, IBT consumption and IBT-PP consumption of household *i* locates.

4.2 Invoices Heading On the basis of these consumption data, additional information is provided including:

- IBT invoice amount T_i^{IBT}, perfectly perceived IBT invoice amount $T_i^{\mathrm{IBT\text{-}PP}}$, and TBSE invoice amount T_i^{TBSE}
- the expenditures associated with captive consumption q_{i0}, with $T_{i,0}^{\mathrm{IBT}} = T\left(q_{i0}\right)$ for IBT and $T_{i,0}^{\mathrm{TBSE}} = T\left(q_{i0}\right)$ for TBSE
- the expenditures associated with variable ("economic") consumptions $q_i^{\mathrm{IBT}} - q_{i,0}$, $q_i^{\mathrm{IBT\text{-}PP}} - q_{i,0}$ and $q_i^{\mathrm{TBSE}} - q_{i,0}$, with $T_i^{\mathrm{IBT}} - T_{i,0}^{\mathrm{IBT}}$, $T_i^{\mathrm{IBT\text{-}PP}} - T_{i,0}^{\mathrm{IBT}}$ and $T_i^{\mathrm{TBSE}} - T_{i,0}^{\mathrm{TBSE}}$ for IBT, IBT-and TBSE (this information is used to calculate some indicators in the field "Quality of the funding")
- the amounts of expenditure borne by the Household to meet its basic needs $\underline{q}_i$ under, respectively, (i) the IBT tested by the user, $\underline{T}_i^{\mathrm{IBT}} = T_{\mathrm{IBT}}(\underline{q}_i)$, and (ii) the TBSE $\underline{T}_i^{\mathrm{TBSE}} = T_{\mathrm{TBSE}}(\underline{q}_i)$ (this information is used to feed the calculation of indcators in the Affordability field)
- the part of IBT, IBT-PP and TBSE invoices linked to non-basic consumption $T_i^{\mathrm{IBT}} - \underline{T}_i^{\mathrm{IBT}}$, $T_i^{\mathrm{IBT\text{-}PP}} - \underline{T}_i^{\mathrm{IBT}}$, and $T_i^{\mathrm{TBSE}} - \underline{T}_i^{\mathrm{TBSE}}$

for, successively, the "EP" drinking water service, the "A" wastewater service and the general (consolidated) "EP/EPA" service. It should be noted that, as with consumption, these expenditure amounts are broken down by consumption blocks:

$T_{i,1}^{\mathrm{IBT}}$, $T_{i,2}^{\mathrm{IBT}}$, …, $T_{i,p}^{\mathrm{IBT}}$

$T_{i,1}^{\mathrm{IBT\text{-}PP}}$, $T_{i,2}^{\mathrm{IBT\text{-}PP}}$, …, $T_{i,p}^{\mathrm{IBT\text{-}PP}}$

…

$\underline{T}_{i,1}^{\mathrm{IBT}}$, $\underline{T}_{i,2}^{\mathrm{IBT}}$, …, $\underline{T}_{i,p}^{\mathrm{IBT}}$

…

with (i) a variable part equal to the sum of these different invoice components :

$$T_{i,v}^{\mathrm{IBT}} = T_{i,1}^{\mathrm{IBT}} + T_{i,2}^{\mathrm{IBT}} + \ldots + T_{i,p}^{\mathrm{IBT}} ,$$

$$T_{i,v}^{\mathrm{IBT\text{-}PP}} = T_{i,1}^{\mathrm{IBT\text{-}PP}} + T_{i,2}^{\mathrm{IBT\text{-}PP}} + \ldots + T_{i,p}^{\mathrm{IBT\text{-}PP}} ,$$

…

$$\underline{T}_{i,v}^{\mathrm{IBT}} = \underline{T}_{i,1}^{\mathrm{IBT}} + \underline{T}_{i,2}^{\mathrm{IBT}} + \ldots + \underline{T}_{i,p}^{\mathrm{IBT}} ,\ \underline{T}_{i,2}^{\mathrm{IBT}} ,\ \ldots,\ \underline{T}_{i,p}^{\mathrm{IBT}}$$

and (ii) a fixed part determined by the subscription amounts (including tax) F_{EP} and $F_{\mathrm{EP}}^{\mathrm{TBSE}}$ (EP service), F_{A} and $F_{\mathrm{A}}^{\mathrm{TBSE}}$ (A service), $F_{\mathrm{EPA}} = F_{\mathrm{EP}} + F_{\mathrm{A}}$ and $F_{\mathrm{EPA}}^{\mathrm{TBSE}} = F_{\mathrm{EP}}^{\mathrm{TBSE}} + F_{\mathrm{A}}^{\mathrm{TBSE}}$ (EPA service), and:

$$F_{\mathrm{IBT}} = \begin{cases} F_{\mathrm{EP}} & \textit{if} \quad \mathrm{ASSAINI} = 0 \\ F_{\mathrm{EPA}} & \textit{if} \quad \mathrm{ASSAINI} = 1 \end{cases} \tag{6.1}$$

$$F_{\mathrm{TBSE}} = \begin{cases} F_{\mathrm{EP}}^{\mathrm{TBSE}} & \textit{if} \quad \mathrm{ASSAINI} = 0 \\ F_{\mathrm{EPA}}^{\mathrm{TBSE}} & \textit{if} \quad \mathrm{ASSAINI} = 1 \end{cases} \tag{6.2}$$

for the general (consolidated) EP / EPA service (as a reminder (see section 3), the variable $\mathrm{ASSAINI}$ is a dummy that takes the value 1 if the household is connected to the collective sanitation network, and 0 otherwise). In addition, each of these variables is broken down into an Operator part (excluding charges and VAT), a Water Agency part (with the collection of charges on the drinking water and wastewater services) and a State part (with the collection of VAT, at different rates on the drinking water and wastewater services).

4.3 "Service Cost" Heading The Cost of service section simply calculates the cost that the household passes on to the service through its consumption. This includes :

(1) a drinking water component :

$$C_i^{\mathrm{EP}} = C_{\mathrm{EP}}\left(q_i\right) = \frac{CF_{\mathrm{EP}}}{n} + c_{\mathrm{EP}} \times q_i \tag{6.3}$$

where (as a reminder) CF_{EP} are the fixed costs of the drinking water service, c_{EP} the cost of producing one cubic metre of drinking water and *n* the number of subscribers to the drinking water service,

(2) a sanitation component A (for households connected to the collective sanitation network):

$$C_i^{\mathrm{A}} = C_{\mathrm{A}}\left(q_i\right) = \frac{CF_{\mathrm{A}}}{n_{\mathrm{A}}} + c_{\mathrm{A}} \times q_i \quad (6.4)$$

with (as a reminder) CF_{A} the fixed costs of "A" service, c_{A} the cost of treating one cubic metre of domestic waste water and n_{A} the number of subscribers to "A" service,

(these values correspond to those entered by the user, at the start of the simulation, with the description of the costs of EP service and A service) and :

(3) a consolidated EP / EPA part with:

$$C_i = C_i^{\mathrm{EP}}$$

for households in Group 1 that are not connected to the public sewerage system and :

$$C_i = C_i^{\mathrm{EPA}} = C_i^{\mathrm{EP}}\left(q_i\right) + C_i^{\mathrm{A}}\left(q_i\right) = \frac{CF_{\mathrm{EP}}}{n} + \frac{CF_{\mathrm{A}}}{n_{\mathrm{A}}} + \left(c_{\mathrm{EP}} + c_{\mathrm{A}}\right) \times q_i \quad (6.5)$$

for households in group 2 that pay for both drinking water and wastewater services.

It should be noted that these service costs are calculated :

(i) excluding taxes and fees (Operator values) ;

(ii) for basic consumption $\underline{q}_i$, captive consumption $q_{i,0}$, IBT consumption q_i^{IBT} , IBT-PP consumption $q_i^{\mathrm{IBT\text{-}PP}}$, and TBSE consumption q_i^{TBSE}

and per balance :

(i) for captive but non-basic consumption $q_{i,0} - \underline{q}_i$ (that is, the water uses for garden maintenance and swimming pool maintenance (excluding user reprocessing)), for the variable ("economic") part of TBSE consumption $q_l^{\mathrm{TBSE}} - q_{i,0}$, for the variable ("economic") part of IBT consumption excluding overconsumption $q_l^{\mathrm{IBT\text{-}PP}} - q_{i,0}$, and for overconsumption $q_l^{\mathrm{IBT}} - q_l^{\mathrm{IBT\text{-}PP}}$.

4.4 Subsidies / Taxes Heading By juxtaposing the various components of the bill with the various components of the "Service Cost", it is identified (for all the Households in the Population module) the various gross and net subsidies and "taxes" that are generated by the IBT which is tested / evaluated by the user. These include:

- the amount of the subsidy/taxation from which the household benefits on the access fee $C_{i0} = \frac{CF}{n} - F$ (see below);
- the amount of the gross subsidies and of the net subsidy the household benefits on its consumption (the net subsidy takes the value 0 when the household is a net contributor to the financing of the service, through its consumption);

- the amounts of gross subsidies and of net subsidies granted to household "DAI";
- the amount of the gross and net contributions to service funding generated on the consumption of household *i* (the net contribution takes the value 0 when the total gross subsidies paid on household consumption is greater than the total revenue collected on this same consumption);
- the amount of the gross and net contributions to service funding which are generated in fine on household *i* (the net contribution takes the value 0 when the household costs in fine more (to the service) than it brings in (through its bill)).

As before, these various financial flows are calculated :

- for each of the services (drinking water, wastewater treatment) and for the general service (drinking water or drinking water and wastewater treatment),
- for the various components of IBT consumption: basic consumption, captive consumption, captive but non-basic consumption, variable part of water consumption, variable part excluding overconsumption, overconsumption, etc.
- for the Operator component and the State component[22]

and, as far as the variable parts are concerned, broken down by consumption block. These household water accounts also provide information on inclusion and exclusion errors, in terms of volume and value, generated by the water tariff calibration set by the user including:

- the volume of basic consumption that is subsidised by the tariff ("true positive"),
- the volume of basic consumption that is taxed/subject to a contribution to finance the service,
- the volume of non-basic consumption $q_i - \underline{q}_i$ which is wrongly subsidised ("false positive"),
- the volume of non-basic consumption on which a margin is generated to finance the service

and:

- the gross and net subsidies granted on basic service (Access Fee and Basic Consumption),
- the gross and net subsidies granted on non-basic service,
- the gross and net "taxes" levied on basic service,
- the gross and net "taxes" levied on non-basic service.

All this information is then used to calculate various indicators in the following fields:

- Incentive effect (for measuring the proper calibration of consumption blocks);
- Equity (with the use of index aiming at measuring the redistributive/anti-redistributive impact of the IBT, which is evaluated/tested by the user),
- Cost recovery (identifying sources of expenditure and funding).

[22] As explained above, given the mechanism of VAT which is an ad valorem tax, the application of an IBT by the operator generates a system of subsidies/taxations form the State on the consumption tax side.

4.5 Other impacts Finally, these water accounts contain information on:

- affordability deficits (amounts of excess spending), as defined by the PAR and CAR, for the IBT and TBSE successively,
- the private costs of mismanagement $T_i^{\mathrm{IBT}} - T_i^{\mathrm{IBT\text{-}PP}}$, linked to tariff misperception, that are borne by the household,
- the contribution to the recovery of environmental costs (based on the values for c_e, c_{EP}, c_{A}, r_{EP}, and r_{A} that are entered by the user, and the connection or not of the household to the collective sanitation network), given the amount of its excise duty charges (which are paid back by the operator to the Water Agency),

as well as on:

- IBT and TBSE household's contributions to the aggregate (social) surplus, in variation from the first-best allocation (see section 9),
- variations in consumer surplus, per subscriber (household) and per capita (members of the family), linked to the implementation of IBT compared to TBSE,

to assess the gains and costs, measured in euros, in social welfare and in private welfare (households, individuals (members of the family)) generated by the implementation of the IBT (which is evaluated/tested by the user).

VII – EVALUATION – AFFORDABILITY

7.1 General information

This field of analysis aims to measure the capacity of the tariff to guarantee access by households/individuals to the public drinking water supply (and public wastewater supply) on reasonable/affordable/sustainable financial terms, at least for their basic needs (understood in the broad sense, i.e. for good living conditions). The issue here is to know to what extent the pricing policy being evaluated/tested by the user meets one of the performance points set by the EU-WFD on water and also targets T1.4 ("By 2030, ensure that all men and women, in particular the poor and the vulnerable, have [...] access to basic services"), T6.1 ("By 2030, achieve universal and equitable access to safe and affordable drinking water for all") and T6.2 ("By 2030, achieve access to adequate and equitable sanitation and hygiene for all [...]") of SDGs 1 "End poverty in all its forms everywhere" and 6 "Ensure access to water and sanitation for all". It should also be emphasized that, considering the principle of equality before the public service (which is one of the characteristics of French regulations on pricing for the EP / EPA service), progressive pricing of the social incentive type, by subsidising water consumption for the first consumption blocks, aims to secure/improve the affordability of the service for all households (universality of service). In this sense, the mechanism differs from the social pricing of the service, where this objective prima facie concerns only poor households.

In order to measure water affordability, the indicators used by the tool are based on the calculation of two conventional ratios that are the CAR [Conventional Affordability Ratio] and the PAR [Potential Affordability Ratio] for each of the households in the Population module. The common idea is to assess the relative effort made by the households to access water and sanitation services (OECD [2010]).

CAR Indicator The CAR relates the household expenditures for water, $T_i = T(q_i)$, to its total income R_i with the calculation of the budget coefficient:

$$\text{CAR}_i = \frac{T(q_i)}{R_i} \quad \text{for } i = 1, \ldots, n \tag{7.1}$$

(see notably Komives et al [2005], Fankhauser & Tepic [2007] and Reynaud [2008]) with an affordability issue that is detected when this ratio exceeds some normatively set threshold. The latter is given by the user with the setting of the primitives of its decision problem (under the heading Social Data in the General Data tab of the MMS). Some international institutions such as the World Bank or the OECD suggest that water bills should not exceed 3-5% of a household's income.

The CAR measure presents a number of limitations, the most important of which is that water consumption linked to uses that do not meet basic needs is included in the calculation of the indicator. In particular, this bias can lead to consider that households that use water wastefully and/or whose water use meets luxury needs are facing an affordability issue, which is clearly unsatisfactory.

PAR Indicator The PAR relates the charge paid by the household for its basic water consumption $\underline{q}_i$, that is $T_{\min,i} = T(\underline{q}_i)$, to the level of its income R_i with the calculation of the following budget coefficient:

$$\text{PAR}_i = \frac{T(\underline{q}_i)}{R_i} \quad \text{for } i = 1, \ldots, n \tag{7.2}$$

(see notably Foster et al. [2006], García-Valiñas et al. [2010] and Martins et al. [2013]). As with the CAR, an affordability issue is detected when this ratio is above a threshold corresponding to the one entered by the user in the same heading as the one entered for the CAR. Tle reference value is usually set at 3%.

The main difficulty with the calculation of PAR is to measure the size of the basic consumption. To do this, one can use standard values proposed by international institutions[23] but it is now accepted, since the work of Howard and Bartram [2003], that the volumes of water that meet basic needs, including essential/life-sustaining needs, while being conditioned by the state of health of individuals and cultural factor (“relationship with water”), environmental factor and technological factor, vary over time, space and from one individual to another. For instance, it is known that pregnant women have higher needs and that children have lower needs for water.

[23] For instance, WHO and UNICEF [2008] suggest that a volume of water of between 15 and 25 l/c/d should be considered as a basic provision to meet domestic needs. Taking collective consumption into account, the World Bank (Komives et al. [2005]) suggests a volume of between 8 and 16 m³ per household per month.

Similarly, in an environment where the temperature is high, it is understood that more water is needed to ensure good living conditions, in a context where social norms regarding the use of water (probably linked to the greater or lesser ease of access to the resource ; see March & Pujol [2009]) also play a role. Lastly, given the induced nature of drinking water consumption (linked to the production of domestic goods and services by the household), the state of available technology matters (through the quality of water-consuming equipment). Ultimately, all these factors mean that basic needs, including essential/vital water needs, are heterogeneous: they vary from region to region and, within a region, from household to household.

Given these factors and following the pioneering work of Barberán & al [2006][24], one possible approach is to measure the distribution of basic needs within a population, living on the area whose the water manager is in charge, using econometric methods[25] . This is precisely what the tool does by making use of the information provided by the econometric model on the determinants of the captive consumption (given the retreatment made by the user). It should be borne in mind, however, that setting a threshold for assessing a family's situation is not without its problems, given the heterogeneity of consumption behaviour. In particular, a poor household that restricts its water consumption may not be identified.

7.2 Presentation of results - aggregate level

The first step is to identify the households (domestic subscribers) facing an affordability issue and the weight of this category of the population in the total population. To this end, the tool calculates the CAR and PAR for each of the households in the Population file. This calculation is performed:

- for the tariff EP / EPA that is evaluated/tested by the user with the PAR IBT variable and the CAR IBT variable,
- for the reference tariff system "TBSE EP / EPA" with the PAR TBSE variable and the CAR TBSE variable.

Next, the values obtained are compared with the thresholds entered by the user (in the Social Data section of the MMS General Data tab) above which the household is considered facing an affordability issue. On this basis, it is created a set of dummy variables:

- two for the CAR ($\mathbf{1}_{\text{CAR}}^{\text{IBT}}$ and $\mathbf{1}_{\text{CAR}}^{\text{TBSE}}$),
- two for the PAR ($\mathbf{1}_{\text{PAR}}^{\text{IBT}}$ and $\mathbf{1}_{\text{PAR}}^{\text{TBSE}}$),

---

[24] The authors specified household water consumption as $q_i = \alpha + \beta N_i$, where α (the collective part of the consumption) and β (the volume required for each additional member) are two parameters to be estimated. Using a dataset of households in Zaragoza (Spain), the fixed consumption was estimated to 3.2 m³ per month and the variable consumption to 2.35 m³ per member.

[25] As previously emphasised, a very common approach in the literature consists in using the linear expenditure system. For mainland France, particular mention should be made of the results of García-Valiñas et al [2009], for whom the average basic consumption (estimated at a high level by captive consumption) is of the order of 108 m³ per household per year, with significant differences between regions (86 m³ for the lowest average value (Brittany) to 169 m³ for the highest average value (Corsica)).

which take the value 1 when the related ratio of household *i* ($\mathrm{CAR}_i^{\mathrm{IBT}}$, $\mathrm{CAR}_i^{\mathrm{TBSE}}$, $\mathrm{PAR}_i^{\mathrm{IBT}}$, $\mathrm{PAR}_i^{\mathrm{TBSE}}$) strictly exceeds the value of the related PAR or CAR-threshold entered by the user, and 0 otherwise. For households facing an affordability issue, the tool then calculates the size of excessive charges, for example :

$$e_i = T(\underline{q}_i) - 0.03R_i$$

for a PAR threshold set at 3%, and displays the value 0 for households that are not. Table 4 below illustrates this treatment of data for calculating the PAR (a similar table appears for the CAR).

i	$\underline{q}$	$T_{\min} = \underline{T}$	R	PAR	$\mathrm{U\_rate}$	$\mathbf{1}_{\mathrm{PAR}}$	e_{PAR}
1	$\underline{q}_1$	$\underline{T}_1$	R_1	$\underline{T}_1 / R_1$	3%	1	$e_1^{\mathrm{PAR}} = \underline{T}_1 - 3\%R_1$
2	$\underline{q}_2$	$\underline{T}_2$	R_2	$\underline{T}_2 / R_2$	3%	0	0
3	$\underline{q}_3$	$\underline{T}_3$	R_3	$\underline{T}_3 / R_3$	3%	1	$e_3^{\mathrm{PAR}} = \underline{T}_3 - 3\%R_3$
$\dots$							
$n-1$	$\underline{q}_{n-1}$	$\underline{T}_{n-1}$	R_{n-1}	$\underline{T}_{n-1} / R_{n-1}$	3%	1	$e_{n-1}^{\mathrm{PAR}} = \underline{T}_{n-1} - 3\%R_{n-1}$
n	$\underline{q}_n$	$\underline{T}_n$	R_n	$\underline{T}_n / R_n$	3%	0	0
Total	$\underline{Q}$					n_{PAR}	Affordability Deficit

Table 4 : PAR unaffordability measurement - basic treatment

On the basis of this information, the tool displays basic descriptive statistics for the 4 variables: PAR IBT, PAR TBSE, CAR IBT and CAR TBSE (see Table 5, page 76), then basic indicators such as the percentage of households facing an affordability issue or the average of over-expenditures for household facing an affordability issue that are presented using a 3Is-type reading grid (Incidence, Intensity, Inequality) applied in the Economics of Poverty (Sen [1976]).

Impact The tool first calculates the mean value and the median value of CARs and PARs over the entire household population (Subscriber Approach) for successively the IBT (which is considered by the user) and the TBSE, which is the reference tariff. On this basis, the tool displays the added value of the IBT, in terms of PAR and CAR, compared with the TBSE, and provides additional information on the distribution of these ratios with (i) the values of the quartiles Q1 and Q3 and of the deciles D1 and D9, (ii) the percentage of the household population that is below the average (the latter makes it easy to calculate the Schutz coefficient given below), (iii) the main dispersion indicators (variance, standard deviation, MAPE, coefficient of variation, interquartile ranges, inter-decile range) to measure the heterogeneity of these ratios (which measure the efforts made by households to meet their basic needs / pay the water bill) within the household population. The value of the Yule coefficient (for measuring symmetry/asymmetry of the distribution) is also provided. On these points, see Table 5, page 76.

	In %		In pourcentage points			
	PAR IBT	PAR TBSE	Delta PAR	CAR IBT	CAR TBSE	Delta CAR
Mean	1.1	3.2	-2.0	2.5	4.1	-1.57
Median	0.7	2.0	-1.3	1.9	2.9	-0.99
Min	0.0	0.1		0.3	0.3	
Max	10.9	21.9		12.6	23.6	
Q1	0.3	0.9		1.0	1.4	
Q3	1.4	3.9		3.3	5.0	
D1	0.1	0.5		0.6	0.8	
D9	2.6	8.0		5.4	9.6	
F (Mean)	68.5	69.3		63.0	68.1	
Variance	1.89	12.64		4.24	15.73	
Standard dev.	1.4	3.6		2.06	3.97	
MAPE	0.9	2.5		1.5	2.9	
Coeff of Variation	1.216	1.119		0.822	0.972	
Interquartile range	1.1	3.0		2.3	3.6	
Interdecile range	2.5	7.5		4.8	8.8	
Yule coefficient	0.29	0.30		0.23	0.18	

Table 5 : Descriptive statistics PAR and CAR -- example

Incidence Next, taking into account the information available in the Population file, it is computed (i) the percentage of households facing an affordability issue with the headcount ratios:

$$H_{\text{CAR}}^{\text{Household}} = \frac{n_{\text{CAR}}}{n} \tag{7.3}$$

$$H_{\text{PAR}}^{\text{Household}} = \frac{n_{\text{PAR}}}{n} \tag{7.4}$$

for, successively, the IBT which is tested by the user and the reference tariff TBSE as well as (ii) the percentage of individuals facing an affordability issue and (iii) the percentage of children facing an affordability issue (for each of these two criteria). The values of these different indicators enable us to assess the incidence of unaffordability (as defined by the PAR and the CAR). See Table 6, page 78, for a numerical illustration.

Intensity The tool first examines the affordability deficit calculated for the population as a whole (thus including households for which the amount of excessive charges is 0), as defined by the CAR and the PAR, with the calculation of the (apparent) average of excessive charges for the IBT, which is tested by the user, and the TBSE, which is the reference tariff. This indicator is a measure in itself of the performance of the tariff policy in terms of affordability. For this reason, the tool displays the average gain linked to the introduction of the IBT and supplements it with data on the median (for which half of the household population has an affordability deficit below the value provided; this value is zero as soon as the Headcount ratio is below 50%). Next, the dispersion/heterogeneity of excessive charges within the domestic subscriber population is measured with the calculation of variances, standard deviations, coefficients de variation and MAPEs. These calculations are then reproduced on the sub-population of household who actually encounter an affordability issue (with, in particular, the calculation of the effective average unaffordability for PAR and CAR criteria). See Table 7 and Table 8, on next page, for a numerical illustration as well as Appendix 2 for additional elements of analysis relating to the interpretation (and the selection) of these indicators.

Inequality Based on the above statistics, it is easy to break down the affordability deficit, as defined by the PAR and CAR criteria, for the household population (subscribers) by:

$$\overline{e}_{\text{PAR}} = \frac{1}{n}\sum_{i=1}^{n} e_i^{\text{PAR}} = \frac{n_{\text{PAR}}}{n} \times \frac{1}{n_{\text{PAR}}}\sum_{i=1}^{n_{\text{PAR}}}\left(T_i - \frac{3}{100}\times R_i\right) = H_{\text{PAR}}^{\text{Household}} \times \overline{e}_{\text{PAR}}^{\text{effective}} \tag{7.5}$$

with, for instance, a threshold value (entered by the user) of 3% for the PAR and in the same way:

$$\overline{e}_{\text{CAR}} = \frac{1}{n}\sum_{i=1}^{n} e_i^{\text{CAR}} = \ldots = H_{\text{CAR}}^{\text{Household}} \times \overline{e}_{\text{CAR}}^{\text{effective}} \tag{7.6}$$

for the CAR. These relationships break down the apparent average of unaffordability (which is in itself an economic policy objective) as the product of the percentage of households facing an unaffordability issue (Incidence) by the effective average of unaffordability (Intensity). Following a suggestion by Sen, the index of this section aim to complete this measure by introducing the third I, that of inequality within the population of households (subscribers) experiencing an affordability issue. For this purpose, the tool gives the values of 2 + 3 basic indicators with:

	In %		In percentage points			
Headcount ratio	PAR IBT	PAR TBSE	Delta PAR	CAR IBT	CAR TBSE	Delta CAR
Household	7.9	32.8	-24.9	28.8	48.3	-19.4
Individuals	7.9	31.6	-23.8	32.6	48.9	-16.3
Children	9.9	35.1	-25.2	38.0	52.3	-14.3

Table 6 : Incidence - headcount ratios

Apparent Deficit	PAR IBT	PAR TBSE	Delta PAR	CAR IBT	CAR TBSE	Delta CAR
Mean	1.39 €	17.47 €	-16.08 €	9.51 €	29.03 €	-19.52 €
Median	0.00 €	0.00 €	0.00 €	0.00 €	0.00 €	
Variance	38.5366	1013.2596		389.6407	1718.7326	
Standard deviation.	6.21 €	31.83 €		19.74 €	41.46 €	
Coeff of Variation	4.47	1.82		2.076	1.428	
MAPE	2.57 €	24.49 €		13.97 €	34.32 €	

Table 7 : Intensity -- apparent affordability gap (in euros per quarter)

Effective Deficit	PAR IBT	PAR TBSE	Delta PAR	CAR IBT	CAR TBSE	Delta CAR
Mean	17.69	53.33	-35.64	33.00	60.17	-27.17
Median	16.15	43.78		27.85	52.73	
D1	2.08	11.03		6.07	9.73	
D9	37.01	102.79		64.50	119.81	
Variance	202.0810	1181.3969		576.8663	1688.4959	
Standard deviation.	14.22	34.37		24.02	41.09	
Coeff of Variation	0.804	0.645		0.728	0.683	
MAPE	11.25	30.19		19.26	35.64	

Table 8 : Intensity -- effective affordability deficit (in euros per quarter)

- the Gini index, which corresponds to the ratio of the area of concentration of unaffordability, defined on the basis of the Lorenz curve of unaffordability, to the area of maximum concentration (describing a situation in which a single household would bear all of the mass of unaffordability measured in the household population) equal itself to ½ (to within $1/n$),
- the Schutz coefficient (or "Robin Hood index"), which gives the percentage of the mass of unaffordability that must be transferred, from household sub-population with an above-average affordability deficit to household sub-population with a below-average affordability deficit, to achieve equality of affordability deficits (measure of the redistributive effort),

for the household population as a whole and for the household population facing an affordability issue, as well as:

- the inter-decile ratio D9/D1, which gives the minimum factor by which the affordability deficit of the first decile (value of the affordability deficit below which 10% of the population is found) must be multiplied to integrate the last decile (value of the affordability deficit above which 10% of the population is found),
- the S90/S10 inter-decime ratio, equal to the ratio of the average affordability deficit for the last decile to the average affordability deficit for the first decile,
- the so-called S80/S20 inter-decime ratio, equal to the ratio of the average affordability deficit for the last quintile to the average affordability deficit for the first quintile of household,

for the sub-population of household facing an affordability issue only (except in extreme cases in which more than 90% (80% in the case of the S80/S20) of the household population would face an affordability issue, the values of the last three indicators involving a division by 0 cannot be calculated for the population as a whole). Ultimately, these elements make it possible to define a synthetic indicator of affordability, à la Sen, the calculation of which is left (at this stage of the tool's development) to the user[26].

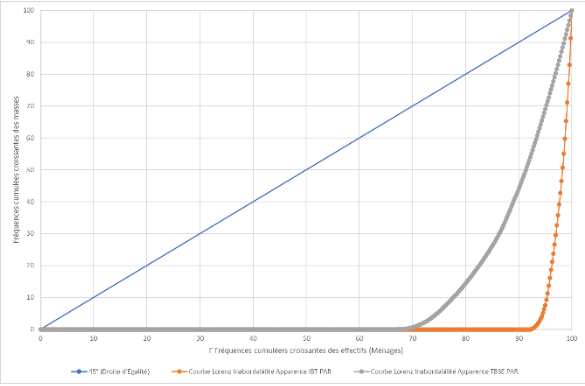
On these different points, see Figure 13, page 80. As usual, Gini index is read by relating two areas:

(1) the first is the concentration area, equal to the area bounded by the 45° line (equality line) and the Lorenz curve $A = L(F)$, with $F = F(x)$ the distribution function of unaffordability x and $A = A(x)$ the share of the mass of unaffordability that goes to the $100F(x)\%$ of the population whose unaffordability is less than x ,

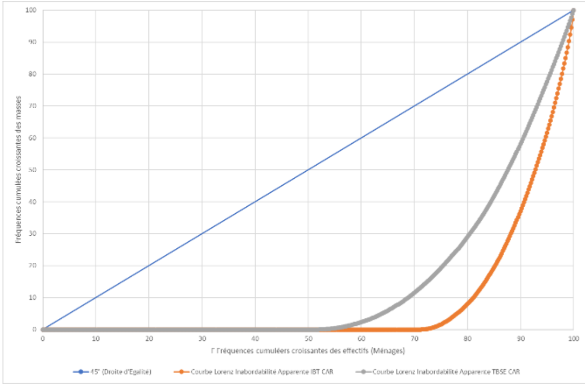
(2) the secund is the maximum concentration area, equal (to within $1/n$) to the area of the OAB triangle (in a situation where the aggregate affordability deficit is concentrated in a single household, the Lorenz curve merges with the abscissa axis for the $100\left(1-\frac{1}{n}\right)\%$ of the population with an affordability deficit equal to 0, before rejoining the point $\left(1,1\right)$ with the introduction of the last household (which bears all the unaffordability) in the Pen's parade of unaffordability).

[26] On this particular point, several options and approaches are possible and may require additional data processing, particularly if one wants to reason in terms of unaffordability per capita (household member). This question will have to be explored in future work.

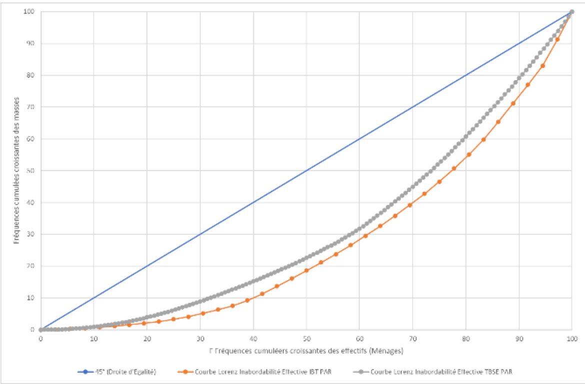
Figure 13 : Lorenz curves for Apparent and Effective Unaffordability deficit, in the sense of CAR and PAR, IBT vs. TBSE



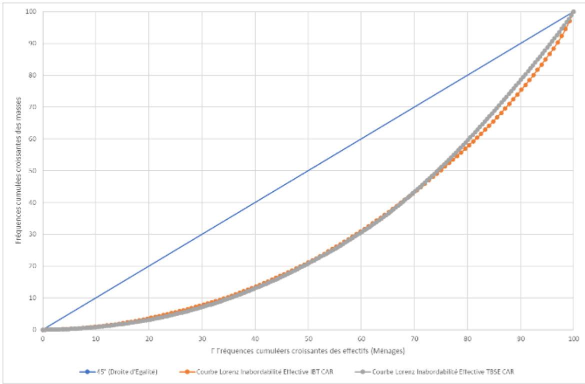
13.1: showing the Lorenz curves for IBT & TBSE PAR Apparent Unaffordability



13.3: showing the Lorenz curves for IBT & TBSE CAR Apparent Unaffordability



13.2: showing the Lorenz curves for IBT & TBSE PAR Effective Unaffordability



13.4: showing the Lorenz curves for IBT & TBSE CAR Effective Unaffordability

The Schutz coefficient corresponds to the maximum vertical distance between the Lorenz curve $A = L(F)$ and the equality line $A = F$. By means of a few calculations, it can be shown that this point (denoted S) on the Lorenz curve (at which the vertical distance $F - L(F)$ is at its greatest) has the coordinates $S = \left(F(\bar{x}), A(\bar{x})\right)$, where $\bar{x}$ is the average level of unaffordability, $F(\bar{x})$ is the percentage of households whose unaffordability is below the average, and $A(\bar{x})$ is the proportion of the total unaffordability that is concentrated in households whose unaffordability is below the average. Combined with the information in Table 7, this last property makes it easy to locate each of the S points.

7.3 Presentation of results - more details

Once these initial elements introduced, the tool provides more detailed information on the distribution of unaffordability within the population by producing tables, based on the CAR and PAR calculations, in which households are broken down into sub-populations according two main criteria (as a first step) with:

- whether households are connected to the collective sanitation network, in which case they are charged the Drinking Water and Sewerage tariff (with the price of water almost doubling), or not connected, in which case they are charged the Drinking Water tariff only;
- whether or not the Household is a poor household, depending on whether its standard of living is below (strictly) or above or equal to the poverty threshold (entered by the user).

It should be noted that, from a methodological point of view, the aim is not only to focus on particular groups of the population but also to break down (where possible) the figures obtained for the general population into its various components calculated for the sub-populations that make it up. To this end, one proceeds as follows.

- Firstly, the tool displays the descriptive statistics calculated for the PAR IBT, PAR TBSE, CAR IBT and CAR TBSE for the sub-population of Households in Group 1 ("Households are not connected to the collective sanitation network / pay for the drinking water service only") and the sub-population of Households in Group 2 ("Households are connected to the collective sanitation network / pay the EPA tariff").

See Table 9, page 82. Next, once these initial observations made, the tool explores:

A) the dimension of **Incidence** by displaying:

- the matrix of Headcount ratios for Households / Individuals / Children, as defined by the PAR and the CAR, for the IBT and TBSE,
- the breakdown of the Household Headcount ratio into its two components Group 1 and Group 2

with for example :

$$H_{\text{PAR}}^{\text{Household}} = \frac{n_1}{n} \times H_{\text{PAR}}^{\text{Household\_G1}} + \frac{n_2}{n} \times H_{\text{PAR}}^{\text{Household\_G2}}$$

	PAR IBT		PAR TBSE	
	G1 (EP)	G2 (EPA)	G1 (EP)	G2 (EPA)
Mean	0.6	1.7	2.0	4.6
Median	0.4	1.2	1.2	3.0
Min	0.0	0.1	0.1	0.2
Max	3.8	10.9	9.9	21.9
Q1	0.2	0.6	0.6	1.7
Q3	0.8	2.3	2.3	6.3
D1	0.1	0.3	0.4	0.9
D9	1.5	4.2	4.7	10.8
F (Mean)	65.5	65.4	65.9	67.7
Variance	0.47	2.92	4.16	18.86
Standard deviation	0.7	1.7	2.0	4.3
MAPE	0.5	1.2	1.5	3.2
Coeff of Variation	1.091	0.991	0.991	1.038
Interquartile range	0.6	1.8	1.7	4.6
Interdecile range	1.3	3.9	4.3	9.9
Yule coefficient	0.38	0.32	0.29	0.42

CAR IBT		CAR TBSE	
G1 (EP)	G2 (EPA)	G1 (EP)	G2 (EPA)
1.8	3.3	2.8	5.59
1.4	2.6	1.9	3.82
0.3	0.3	0.3	0.3
8.6	12.6	12.8	23.6
0.7	1.5	1.0	2.4
2.5	4.5	3.5	7.6
0.5	1.0	0.7	1.3
3.9	6.3	6.3	12.3
62.0	61.0	48.3	65.0
2.18	5.52	6.43	22.50
1.48	2.35	2.54	4.7
1.1	1.8	1.9	3.6
0.804	0.714	0.906	0.849
1.7	3.0	2.5	5.2
3.4	5.3	5.7	11.0
0.20	0.26	0.31	0.44

Table 9 : Descriptive statistics PAR and CAR -- Focus on Group 1 ("EP" customer segment) vs. Group 2 ("EPA" customer segment)

in the case of the Household PAR (the same formula applies for the CAR). See Table 10, page 84, for a numerical illustration for PAR IBT and PAR TBSE (a similar table is displayed for CAR IBT and CAR TBSE).

B) the **Intensity** dimension by displaying in dedicated tables :

(1) the conditional averages of affordability deficits, apparent and effective, of Groups 1 and 2, as defined by PAR and CAR, for the IBT (which is tested/assessed by the user) and the TBSE ;

(2) the breakdown of the affordability deficit by group ;

(3) the conditional variances of the affordability deficits listed in (1) ;

(4) the decomposition by groups of the variance of the related affordability deficit into its intra-group component (equal to the average of the conditional variances) and its inter-group component (variance of the conditional averages);

(5) calculation of the associated correlation ratio, which measures the proportion of individual heterogeneity (as measured by variance) attributable to inter-group differences

with, for example :

$$\overline{e}_{\text{PAR}}^{\text{G1}} = \frac{1}{n_1} \times \sum_{i=1}^{n_1} e_{i,\text{PAR}} = \frac{1}{n_1} \times \sum_{i=1}^{n_1} \max\left[\underline{T}_i - \frac{3R_i}{100}, 0 \right]$$

$$\overline{e}_{\text{PAR}}^{\text{G2}} = \frac{1}{n_2} \times \sum_{i=1}^{n_2} e_{i,\text{PAR}} = \frac{1}{n_2} \times \sum_{i=1}^{n_2} \max\left[\underline{T}_i - \frac{3R_i}{100}, 0 \right]$$

$$\overline{e}_{\text{PAR}} = \frac{1}{n} \times \sum_{i=1}^{n} e_{i,\text{PAR}} = \frac{n_1}{n} \times \left(\frac{1}{n_1} \times \sum_{i=1}^{n_1} e_{i,\text{PAR}}^{\text{G1}} \right) + \frac{n_2}{n} \times \left(\frac{1}{n_2} \times \sum_{i=1}^{n_2} e_{i,\text{PAR}}^{\text{G2}} \right) = \frac{n_1}{n} \times \overline{e}_{\text{PAR}}^{\text{G1}} + \frac{n_2}{n} \times \overline{e}_{\text{PAR}}^{\text{G2}}$$

$$V_1 = V\left(e_{\text{PAR}}^{\text{G1}}\right) = \frac{1}{n_1} \times \sum_{i=1}^{n_1} \left(e_{i,\text{PAR}} - \overline{e}_{\text{PAR}}^{\text{G1}} \right)^2$$

$$V_2 = V\left(e_{\text{PAR}}^{\text{G2}}\right) = \frac{1}{n_2} \times \sum_{i=1}^{n_2} \left(e_{i,\text{PAR}} - \overline{e}_{\text{PAR}}^{\text{G2}} \right)^2$$

$$V\left(e_{\text{PAR}}\right) = V_{\text{intra}}\left(e_{\text{PAR}}\right) + V_{\text{inter}}\left(e_{\text{PAR}}\right)$$

$$V_{\text{intra}}\left(e_{\text{PAR}}\right) = \frac{n_1}{n} \times V\left(e_{\text{PAR}}^{\text{G1}}\right) + \frac{n_2}{n} \times V\left(e_{\text{PAR}}^{\text{G2}}\right)$$

$$V_{\text{inter}}\left(e_{\text{PAR}}\right) = \frac{n_1}{n} \times \left(\overline{e}_{\text{PAR}}^{\text{G1}} - \overline{e}_{\text{PAR}} \right)^2 + \frac{n_2}{n} \times \left(\overline{e}_{\text{PAR}}^{\text{G2}} - \overline{e}_{\text{PAR}} \right)^2$$

$$\eta_{e_{\text{PAR}}}^2 = \frac{V_{\text{inter}}\left(e_{\text{PAR}}\right)}{V\left(e_{\text{PAR}}\right)}$$

in the case of the Household PAR (the same set of formulas applies for the CAR). See Table 11, page 84.

Table 10 : Incidence (Headecount ratios) – Focus and Breakdown Group 1 ("EP" Service) vs. Group 2 ("EPA" Service)

		PAR IBT			PAR TBSE			**Delta**		
	f_j (in %)	Household	Individuals	Children	Household	Individuals	Children	Household	Individuals	Children
"EP" (A unconnected)	54.1	**1.2**	1.8	1.8	18.1	19.0	20.9	-16.9	-17.2	-19.1
"EPA" (A connected)	45.9	**15.7**	15.8	19.2	50.0	46.6	51.3	-34.3	-30.8	-32.1
Total Population	100.0	**7.9**	7.9	9.9	32.8	31.6	35.1	-24.9	-23.8	-25.2

Table 11 : I ntensity (apparent and actual affordability deficits) -- Focus and Breakdown Group 1 ("EP" Service) vs. Group 2 ("EPA" Service)

Apparent Deficit	**Household**	**PAR IBT**				**PAR TBSE**				
	f_j (in %)	Mean	Variance			f_i (en %)	Mean	Variance		
"EP" (A unconnected)	54.1	0.06	0.3437	**Between**	2.0827	54.1	4.50	126.2414	**Between**	198.3668
"EPA" (A connected)	45.9	2.96	79.0985	**Within**	36.4540	45.9	32.77	1628.1573	**Within**	814.8928
Total Population	100.0	1.39	38.5366	**Corr. Coeff**	5.4	100.0	17.47	1013.2596	**Corr. Coeff**	19.6

Effective Deficit		**PAR IBT**				**PAR TBSE**				
	g_j (in %)	Mean	Variance			g_j (en %)	Mean	Variance		
"EP" (A unconnected)	8.3	5.14	2.3430	**Between**	14.3157	**30.0**	24.83	191.2013	**Between**	348.1182
"EPA" (A connected)	91.7	18.83	204.6219	**Within**	187.7653	**70.0**	65.54	1108.4547	**Within**	833.2787
Total Population	100.0	17.69	202.0810	**Corr. Coeff**	7.1	100	53.33	1181.3969	**Corr. Coeff**	29.5

C) the **Inequality** dimension with the decomposition by groups of the Gini index, applying the methodology of Lambert & Aronson [1993] and that of Dagum [1997], into its three components Inter, Intra and Transvariation (and the provision of additional information) for, again, the apparent and effective affordability deficits, in the sense of the PAR and the CAR, for the IBT (which is tested / evaluated by the user) and the TBSE.

See (i) Table 12, page 86, for a numerical illustration with the breakdown by groups, G1 vs. G2, of the Gini index of apparent unaffordability (including 0) as defined by the PAR, and (ii) Appendix 3 for a description of the methodology.

These focuses and decompositions are then reproduced for, on the one hand, a breakdown of the population into poor households and non-poor households and, on the other hand, by cross-referencing this latter differentiation with the EP vs. EPA distinction to produce contingency tables providing information on the distribution of unaffordability, its extent and the impact of the IBT compared with the TBSE with tables of the form :

	Poor	Non poor	Ensemble
G1 (service EP)	$H_{1,\text{Poor}}$	$H_{1,\text{Non\_Poor}}$	H_1
G2 (service EPA)	$H_{2,\text{Poor}}$	$H_{2,\text{Non\_Poor}}$	H_2
Total Population	H_{Poor}	$H_{\text{Non\_Poor}}$	H

Heacount ratios matrix

	Poor	Non poor	Ensemble
G1 (service EP)	$\overline{e}_{1,\text{Poor}}$	$\overline{e}_{1,\text{Non\_Poor}}$	$\overline{e}_1$
G2 (service EPA)	$\overline{e}_{2,\text{Poor}}$	$\overline{e}_{2,\text{Non\_Poor}}$	$\overline{e}_2$
Total Population	$\overline{e}_{\text{Poor}}$	$\overline{e}_{\text{Non\_Poor}}$	$\overline{e}$

Apparent Affordability Deficit (Average)

	Poor	Non poor	Ensemble
G1 (service EP)	$\overline{e}^*_{1,\text{Poor}}$	$\overline{e}^*_{1,\text{Non\_Poor}}$	$\overline{e}^*_1$
G2 (service EPA)	$\overline{e}^*_{2,\text{Poor}}$	$\overline{e}^*_{2,\text{Non\_Poor}}$	$\overline{e}^*_2$
Total Population	$\overline{e}^*_{\text{Poor}}$	$\overline{e}^*_{\text{Non\_Poor}}$	$\overline{e}^*$

Effective Affordability Deficit (Average)

	Poor	Non poor	Ensemble
G1 (service EP)	$f_{1,\text{Poor}}$	$f_{1,\text{Non\_Poor}}$	f_1
G2 (service EPA)	$f_{2,\text{Poor}}$	$f_{2,\text{Non\_Poor}}$	f_2
Total Population	f_{Poor}	$f_{\text{Non\_Poor}}$	100

Household composition Effective affordability (%)

to assess the social aspect of the tariffs, i.e. their ability to support poor households in terms of affordability.

Table 12 : Inequality - Focus and breakdown Group 1 ("EP" Service) vs. Group 2 ("EPA" Service) of the Gini index of apparent unaffordability, as defined by the PAR

Gini PAR IBT		En %
Between	51.7	54.1
Within	42.0	43.9
Transvariat.	1.9	2.0
Total Population	95.6	100.0

Gini PAR TBSE		En %
Between	40.2	50.7
Within	32.1	40.5
Transvariat	7.0	8.8
Total Pop	79.3	100.0

Breakdown intergroup Gii index

Gini Index matrix (en %)				
	G1	G2		
G1	99.0	99.4		3.6
G2	99.4	90.9	95.9	

	G1	G2		
G1	87.7	89.0		13.2
G2	89.0	64.6	75.8	

Weight matrix - calculation				
	f_j	α_j	w_{ij}	
G1	54.1	2.4	0.013	0.539
G2	45.9	97.6	0.539	0.447
	100.0	100.0		

	f_j	alpha_j	w_ij	
G1	54.1	14.0	0.076	0.530
G2	45.9	86.0	0.530	0.394
	100.0	100.0		

Contributions matrix (in perrcentage points)		
	G1	G2
G1	1.3	1.9
G2	51.7	40.7
		95.6

	G1	G2
G1	6.7	7.0
G2	40.2	25.5
		79.4

	IBT	TBSE
Economid Distance 2-1	0.9642	0.8516
Overlap ratio 1-2	0.0358	0.1484

VIII – EVALUATION - THE INCENTIVE EFFECT

This field of analysis aims to measure the ability of the tariff to induce household to use the resource sparingly. In addition to the mandatory nature of this moment in the evaluation of the socio-economic performance of the water (and wastewater) pricing policy (which is one of the performance points listed by the EU-WFD), it should be remembered that an increasing block tariff of the social incentive type also aims, by setting high prices for levels of consumption considered to be high, to induce large consumers to reduce their consumption. In this respect, it is spontaneously necessary and coherent to see to what extent IBT actually satisfies this objective.

The tool then provides information on 3 major items relating successively to (*i*) the impact on consumption of the structuring of the unit price scale (compared with a structurally balanced two-part tariff) and the extent of the over-consumption that comes with, linked to a poor perception of the tariff, (*ii*) the costs of mismanagement that this poor tariff perception imposes on household, and (*iii*) the proper calibration of consumption blocks with the measurement of inclusion errors and exclusion errors in volume.

8.1 Impact on consumption

The tool first displays:

- the average household consumption in relation to the IBT, which is evaluated/tested by the user, making use of the econometric model defined in the Demand module (by the user with the setting of the values for the response coefficients) to calculate the predicted consumption of each household in the Household Subscriber file:

$$\bar{q}_{\text{IBT}} = \frac{1}{n}\sum_{i=1}^{n} q^d\left(q_{0i}, R_i, F, \pi_1, \ldots, \pi_p, k_1, \ldots, k_{p-1} \middle| \kappa_0\right) \quad (8.1)$$

with q_{0i} the (estimated) captive consumption of household *i* and $\kappa_0 \in [0,1]$ the value of the perception parameter entered by the user (in the Demand module)[27];

- the average household consumption in relation to the IBT when the latter is properly perceived:

$$\bar{q}_{\text{IBT-PP}} = \frac{1}{n}\sum_{i=1}^{n} q^d\left(q_{0i}, R_i, F, \pi_1, \ldots, \pi_p, k_1, \ldots, k_{p-1} \middle| 1\right) \quad (8.2)$$

- the average household consumption of TBSE $\left(F_{\text{TBSE}}, \pi_{\text{TBSE}}\right) = \left(\frac{CF}{n}, c\right)$:

$$\bar{q}_{\text{TBSE}} = \frac{1}{n}\sum_{i=1}^{n} q^d\left(q_{0i}, R_i, \frac{CF}{n}, c\right) \quad (8.3)$$

[27] Pricing parameters are inclusive of VAT.

	In cubic metre per quarter				
	IBT	IBT_PP	TBSE	Δ_1	Δ_2
Mean	35.8	31.7	47.5	-11.7	-15.8
Median	34.1	30.0	45.3	-11.1	-15.3
Min	11.1	10.0	9.9		
Max	69.5	64.5	110.5		
Q1	29.6	26.8	35.2		
Q3	41.3	35.6	57.6		
D1	22.8	20.0	26.5		
D9	49.8	44.1	71.0		
F (Mean)	58.8	65.1	56.1		
Variance	110.25	88.22	314.77		
Standard deviation	10.5	9.4	17.7		
MAPE	7.9	6.7	13.7		
Coeff of Variation	0.294	0.297	0.374		
Interquartile range	11.7	8.8	22.3		
Interdecile range	27.0	24.1	44.5		
Yule Coefficient	0.23	0.28	0.10		
Gini Index	0.160	0.157	0.207		
Schutz coefficient	0.110	0.106	0.144		
Interdecile ratio	2.184	2.205	2.683		
Interdecime ratio	3.016	3.028	4.020		
S80/S20	2.301	2.298	2.928		

Table 13 : IBT, IBT-PP and TBSE consumption (in cubic metres per quarter) - Main Descriptive Statistics

(based on the information provided by the user on production and distribution costs), which is the reference tariff used to measure the gains/losses of the IBT in this field of analysis (similar to what is performed for affordability). The classification of the 3 situations TBSE, IBT and IBT with perfect perception (IBT-PP) on the basis of average consumption values in itself makes it possible to assess the potential and actual incentive nature of the increasing block tariff which is evaluated/tested by the user.

These three averages are then supplemented, as for the affordability field, by the calculation of other basic statistics (see Table 13) with:

- the minimum and maximum values, the first and third quartiles, the first and last decile, and the percentage of households whose consumption is below average, to get the main information about the profile of the statistical series (these different elements are used to construct the box plots of the distribution of the variable of interest) and:

- variance, standard deviation, coefficient of variation, MAPE, interquartile range, interdecile range

to measure the dispersion of household water consumption (and asymmetry with the Yule coefficient) and:

- Gini index, inter-decile ratio, inter-decimal ratio, inter-quintime ratio and Schutz coefficient

to measure the concentration of household water consumption and determine whether there is a fringe of large consumers. All of information is provided for the IBT, the IBT with perfect tariff perception and the TBSE.

It should be noted that overall consumption IBT, $Q_{\mathrm{IBT}} = n \times \bar{q}_{\mathrm{IBT}}$, can be higher or lower than overall consumption TBSE, $Q_{\mathrm{TBSE}} = n \times \bar{q}_{\mathrm{TBSE}}$, for the following two reasons:

- At first, the structuring of the unit price scale around the unit variable cost *c* means that a progressive tariff subsidises "small" consumers (who will then increase their consumption, compared with the TBSE) and taxes "large" consumers (who will in turn reduce their consumption, compared with TBSE), with a net effect on overall consumption that will depend on the calibration of the IBT (with, in particular, the size of the consumption blocks that are subsidised and the subsidy rates $\sigma_j = \pi_j - c$ that are granted in each of the subsidised consumption blocks)[28, 29];
- Secundly, the introduction of block pricing (which is somewhat complex) generates over-consumption due to poor perception of the tariff (with underestimation of the marginal price/marginal cost of consumption) as soon as the perception parameter κ (entered by the user) is less than unity.

This over-consumption, which concerns all households in block 2 and above, therefore pushes up the average IBT consumption, compared with a situation where the tariff would be perfectly perceived (case $\kappa = 1$)[30] . To deal with the first topic mentioned above, the tool supplements this information with:

- calculation of the average variation in consumption when it is increasing, as well as the percentage of households whose consumption is increasing
- calculation of the average variation in consumption when it is decreasing, as well as the percentage of households whose consumption is decreasing

for IBT and IBT-PP, compared with TBSE, in two dedicated tables:

[28] These properties have been corroborated empirically. *Cf.* in particular Mayol [2017].

[29] *See* also Ito [2014] for a similar discussion of the effects of electricity pricing with Californian data.

[30] As a reminder, the specification of the econometric demand model estimated by BCP, with an average price which is calculated excluding the amount of the subscription, means that there is no poor perception of the tariff for households whose consumption places them in block 1, and this good perception of the marginal cost of consumption also applies to any two-part tariffs (F, π) including therefore the TBSE for which $(F, \pi) = \left(\frac{CF}{n}, c\right)$.

In pourcentage		IBT_PP		
		Upward	Downward	
IBT	Upward	0.9	0.8	1.7
	Downward	***	98.3	98.3
		0.9	99.1	100.0

Breakdown of subscriber population (%)

In cubic metre		IBT_PP		
		Upward	Downward	
IBT	Upward			19.0
	Downward			36.1
		14.3	31.8	

Conditional average variation in Household consumption (in cubic metres per quarter)

Table 14 : Descriptive statistics for "small" consumers vs. "large" consumers

Besides, to deal with the secund topic mentioned above, the tool also displays statistics relating to overconsumption, in particular :

- calculation of the average, taken from households, of excess consumption
- calculation of the average, taken from households, of excess consumption per capita
- calculation of the average, based on individuals, of excess consumption per capita

for the population of households and the population of individual family members that over-consume (so-called actual averages), i.e. for households whose IBT consumption is located in block 2 or above. Voir Table 15, page 91.

Once this information of level 1 made available, the tool displays (as for affordability) an information of level 2 with focuses and breakdowns by groups of some of these indicators with:

- the breakdown of average IBT, IBT-PP and TBSE consumption
- the breakdown of the variance in IBT, IBT-PP and TBSE consumption
- the breakdown of the Gini index for IBT, IBT-PP and TBSE consumption

into its various components to successively:

- households broken down into two sub-populations according to (i) whether they face the "EP tariff" alone (Group 1) or the "EPA" tariff (Group 2),
- households broken down into two sub-populations depending on whether they are part of the most deprived households (group of poor households) or not (group of non-poor households).

Overconsumptions			
% of over-consumers	Household	Per Capita (Household)	Individuals
	93.7		97.0
Mean	6.1	2.1	1.9
Median	6.0	1.9	1.7
Min	0.3	0.2	0.2
Max	13.3	6.4	6.4
Q1	5.1	1.4	1.3
Q3	7.3	2.6	2.2
D1	4.2	1.1	1.1
D9	7.8	3.7	3.0
F (Mean)	51.4	60.2	58.4
Variance	3.83	1.09	0.70
Standard dev.	2.0	1.0	0.8
MAPE	1.4	0.8	0.6
Coeff of Variation	0.321	0.487	0.452
Interquartile range	2.2	1.2	0.9
Interdecile range	3.6	2.6	1.9
Yule coefficient	0.18	0.12	0.10
Gini index	0.167	***	0.238
Schutz coefficient	0.114	***	0.168
Interdecile ratio	1.866	***	2.808
Interdecime ratio	3.801	***	4.561
S80/S20	2.417	***	3.282

Table 15 : Descriptive statistics - Effective overconsumption (in cubic meter per quarter)

The tool then supplements these figures with the production of contingency tables giving information on:

- the distribution of households whose consumption is increasing
- the distribution of households whose consumption is decreasing
- the breakdown of increases in consumption
- the breakdown of decreases in consumption

for IBT and IBT-PP versus TBSE. In a similar way to affordability, the tool then breaks down into groups:

- the percentage of household over-consumers
- the effective average excess consumption per subscriber (household) and per capita (family member)

- the variance of the effective excess consumption per subscriber (household) and per capita (family member)
- the Gini index of the effective excess consumption per subscriber (household) and per capita (family member),

next suplements this information by producing contingency tables:

	Poor	Non poor	Ensemble
G1 (EP service)			
G2 (EPA service)			
Total Population			

giving (i) the joint distribution of overconsumers and (ii) the joint distribution of overconsumption for the household population (domestic subscribers) and the individual population (family members). These elements conclude the analysis of the impact of the IBT on household water consumption.

8.2 The costs of mismanagement

Knowing the excess consumption (linked to poor perception of the tariff) for each household in the Population file, the tool supplements this information by calculating the excess expenditure generated by this poor perception, for each household in the Population file, with the difference between the amount of two invoices: the amount of water expenditure actually paid minus the amount of expenditure that would be paid for a correctly perceived IBT consumption:

$$\Delta T_i = T\left(q^d\left(q_{0i}, R_i, F, \pi_1, \ldots, \pi_p, k_1, \ldots, k_{p-1} | \kappa_0\right)\right) - T\left(q^d\left(q_{0i}, R_i, F, \pi_1, \ldots, \pi_p, k_1, \ldots, k_{p-1} | 1\right)\right) \qquad (8.4)$$

as well as over-spending per capita (family member):

$$\frac{\Delta T_i}{N_i} = \frac{T\left(q^d\left(q_{0i}, R_i, F, \pi_1, \ldots, \pi_p, k_1, \ldots, k_{p-1} | \kappa_0\right)\right)}{N_i} - \frac{T\left(q^d\left(q_{0i}, R_i, F, \pi_1, \ldots, \pi_p, k_1, \ldots, k_{p-1} | 1\right)\right)}{N_i} \qquad (8.5)$$

for $i = 1, \ldots, n$.

These quantities give an estimate of the cost of poor management that is borne privately by the household (allocative inefficiency) on the basis of a quarter of consumption (billing period). This information can be used:

- to assess the financial returns, which are generally quite low, for households of a capacity investment aimed at understanding the nature of the economic calculation to be implemented for an optimal management of their domestic uses

(the latter can be compared with the cognitive costs that seem to be relatively high when nonlinear incentive schemes are at stake; see notably Bhargava et al [2017], Grubb & Osborne [2015], Chetty & Saez [2015] and Cartter & Milon [2005]), or even:

	In euros per quarter				
	IBT	IBT_PP	TBSE	Δ_1	Δ_2
Mean	80.69	62.99	137.07	-56.38	-74.1
Median	70.87	58.73	127.97	-57.11	-69.2

Min	24.02	24.02	59.41
Max	323.80	230.44	281.83
Q1	52.34	38.00	97.48
Q3	99.65	76.55	175.84
D1	38.50	33.44	83.96
D9	130.04	92.28	194.24
F (Mean	62.0	63.5	51.8

Variance	1792.3861	956.2686	2061.6866
Standard Deviation	42.34	30.92	45.41
MAPE	31.13	21.67	39.90
Coeff of Variation	0.525	0.491	0.331

Interquartile range	47.31	38.55	78.36
Interdecile range	91.54	58.84	110.28
Yule coefficient	0.22	-0.08	0.22

Gini index	0.267	0.244	0.188
Schutz coefficient	0.193	0.172	0.146
Interdecile ratio	3.4	2.8	2.3
Interdecime ratio	5.2	4.4	2.8
S80/S20	3.7	3.3	2.4

Table 16 : Distribution of IBT, IBT-PP and TBSE invoices - Main descriptive statistics

Mismanagement cost (in euros per quarter)			
	Household	Per Capita (Household)	Individuals
% of over-consumers	93.7		97.0
Mean	18.90	6.04	5.76
Median	15.15	5.17	4.97
Min	1.03	0.55	0.55
Max	93.37	23.34	23.34
Q1	10.62	3.55	3.51
Q3	22.44	7.55	7.43
D1	5.06	2.35	2.30
D9	35.41	10.63	9.96
F (Mean)	56.9	57.7	58.6
Variance	170.8909	14.1313	11.8767
Standard deviation	13.07	3.8	3.45
MAPE	9.34	2.8	2.61
Coeff of Variation	0.692	0.622	0.598
Interquartile range	11.83	4.0	3.9
Interdecile range	30.35	8.3	7.7
Yule coefficient	0.23	0.19	0.25
Gini index	0.352	***	0.313
Schutz coefficient	0.247	***	0.226
Interdecile ratio	7.0	***	4.3
Interdecime ratio	12.4	***	8.0
S80/S20	7.8	***	5.2

Table 17 : Effective mismanagement cost (per household and per capital) – Main descriptive statistics

- for the calculation of the social returns of nudges or boosts targeting the proper understanding of the increasing block tariff system (with, in the first case, the provision of information on the marginal cost of consumption like in Jessoe K. & Raspon D. [2014], in the second the modification of the perception parameter κ like in Brick et al. {2017]).

The model then provides basic descriptive statistics to assess the extent and distribution of effective mismanagement costs within the population (including mean, median, variance, standard deviation, coefficient of variation and Gini index; see Table 17, page 94). As with over-consumption, this information is also provided by subgroups with households broken down into 2 sub-populations:

- depending on whether they are connected to the sewerage network, in which case they face the "EPA" tariff, or whether they are not connected, in which case they face the EP tariff only;
- according to whether or not they belong to the group of poor households (based on the value of the poverty threshold entered by the user).

As the distribution of over-consuming households obtained by crossing these two criteria has already been provided at this stage, this information is supplemented with a contingency table specifying the joint distribution of these mismanagement costs. Besides, the main descriptive statistics on the distribution of invoices including VAT, in euros per quarter, for IBT, IBT PP and TBSE are given at the very start of the analysis, similar to that given for consumption. See Table 16, page 93.

8.3 Assessing the proper calibration of the tariff

By setting up a system of subsidies and "taxes" based on consumption, Increasing Block Tariff acts as a classifier. In particular, it is considered:

(i) in the subsidised consumption blocks, that the units consumed are basic units;

(ii) in the consumption blocks that are not subsidised, that the units consumed are not basic units.

The tool then implements a strict approach[31] whereby it is considered:

(iii) that a cubic metre / a unit of basic consumption $\underline{q}$ (basic unit) that is not subsidised is an issue (exclusion-type inefficiency),

(iv) a cubic metre / a unit of non-basic consumption $q-\underline{q}$ that is subsidised is also an issue (inclusion-type inefficiency).

These inefficiencies are then measured with standard indicators, which are used in the fields of medicine (biostatistics), information technology (automatic classification) and evaluation of public policies (aid programmes in particular). The inefficiencies measured here relate to errors in volume, while errors in value (linked to the taxation of basic consumption and the subsidisation of non-basic consumption) are examined using appropriate indicators in the fields of Equity and Cost Recovery (with the "Quality of the Funding" topic).

[31] When the IBT (considered by the user) specifies a unit price $\pi_j = c$ (sale at cost in block *j*), the fact of not being subsidised is not equivalent to being taxed (margined) and there is a room for adopting a non-strict approach.

8.3.1 The confusion matrix

As pointed out above, an IBT acts (de facto) as a classifier and, when a "strict" approach is adopted, the confusion matrix originally given by:

		True value	
		D_+	D_-
Predicted Value	T_+	True positive	False positive
	T_-	False negative	True negative

Table 18 : Confusion matrix

(T for Test, D for Disease, D_+: "the disease is present", D_- : "it is not present", T_+ : "the Test result is positive/concludes the disease is present", T_- : "the Test result is negative/concludes the disease is not present) takes the following form:

		True value	
		D_+ = basic	D_- = non-basic
Treatment	T_+ = subsidies	True positive	False positive
	T_- = does not subsidy	False negative	True negative

Table 19 : Confusion matrix II

(T for Treatment, T_+ : "the unit is subsidised", T_- : "the unit is not subsidised", D_+ : "the unit is a basic unit", D_- : "the unit is not a basic unit"). Within this framework[32], the contingency table associated with the matrix "Table 19":

	D_+	D_-	
T_+	n_{++}	n_{+-}	$\|T_+\|$
T_-	n_{-+}	n_{--}	$\|T_-\|$
	$\|D_+\|$	$\|D_-\|$	n

Table 20 : Contingency table

with:

- n_{++} the number of true positives (instances rightly processed),
- n_{+-} the number of false positives (instances wrongly processed),
- n_{-+} the number of false negatives (instances wrongly not processed)
- n_{--} the number of true negatives (instances rightly not processed),

becomes:

[32] It should be emphasized that the population studied here are not households but cubic metres that is, units of service consumed by household subscribers.

	D_+	D_-	
T_+	$\underline{Q}^+$	$(Q-\underline{Q})^+$	$\lvert Q^+\rvert$
T_-	$\underline{Q}^-$	$(Q-\underline{Q})^-$	$\lvert Q^-\rvert$
	$\underline{Q}$	$Q-\underline{Q}$	Q

Table 21 : Contingency table II

(this confusion matrix is calcuted and displayed by the tool) with:

- $Q=\sum_{i=1}^{n} q_i$ the overall consumption (service level);
- $\lvert D_+\rvert = \underline{Q} = \sum_{i=1}^{n} \underline{q}_i$ the overall basic consumption (basic service level);
- $\lvert D_-\rvert = Q-\underline{Q} = \sum_{i=1}^{n}\left(q_i-\underline{q}_i\right)$ the overall non-basic consumption (non-basic service level);
- $Q^+ = \sum_{i=1}^{n} q_i^+$ the overall consumption which is subsidised;
- $Q^- = \sum_{i=1}^{n} q_i^-$ the overall consumption which is not subsidised;

(unit m³) and:

$$n_{++} = \sum_{i=1}^{n} \underline{q}_i^+ = \underline{Q}^+$$

$$n_{+-} = \sum_{i=1}^{n} \left(q_i-\underline{q}_i\right)^+ = \left(Q-\underline{Q}\right)^+$$

$$n_{-+} = \sum_{i=1}^{n} \underline{q}_i^- = \underline{Q}^-$$

$$n_{--} = \sum_{i=1}^{n} \left(q_i-\underline{q}_i\right)^- = \left(Q-\underline{Q}\right)^-$$

with:

- $\underline{q}_i^+$ the basic consumption of household *i* which is subsidised (treated), $\underline{Q}^+$ the total basic consumption which is subsidised;
- $\left(q_i-\underline{q}_i\right)^+$ the subsidised non-basic consumption of household *i*, $\left(Q-\underline{Q}\right)^+$ the total subsidised non-basic consumption;
- $\underline{q}_i^-$ the basic consumption of household *i* that is not subsidised, $\underline{Q}^-$ the total basic consumption that is not subsidised;
- $\left(q_i-\underline{q}_i\right)^-$ the non-basic consumption of household *i* that is not subsidised, $\left(Q-\underline{Q}\right)^-$ the total non-basic consumption that is not subsidised.

8.3.2 Basic and other indicators

Next and based on the information provided by Table 21, page 97, the tool calculates the values taken by various indicators starting with 4 key-ones (and their complements) with:

- Sensitivity (True Positive Rate) and Miss rate (False Negative Rate),
- Specificity (True Negative Rate) and Anti-specificity (False Positive Rate),
- Positive Predictive Value (Accuracy) and False Discovery Rate,
- Negative Predicted Value and False Omission Rate.

This information (with also the prevalence $p = \underline{Q}/Q$ and the subsidy rate $s = Q_+/Q$) is displayed in a specific table (see Table 22, page 99), next shown on a specific diagram (see Figure 14 below) called the ROC Space (Receiver Operating Characteristic) that enables to compare the performance of different IBTs (as classifiers). Last, the tool concludes the assessment with the calculation of other indicators commonly used to gauge the quality of a classifier with:

- Accuracy (ACC) and adjusted accuracy (ACC*),
- Youden J,
- Kappa Score and Cohen ratio,
- Jaccard Index.

See Appendix 4 for a detailed presentation of these different indicators and the ROC space.

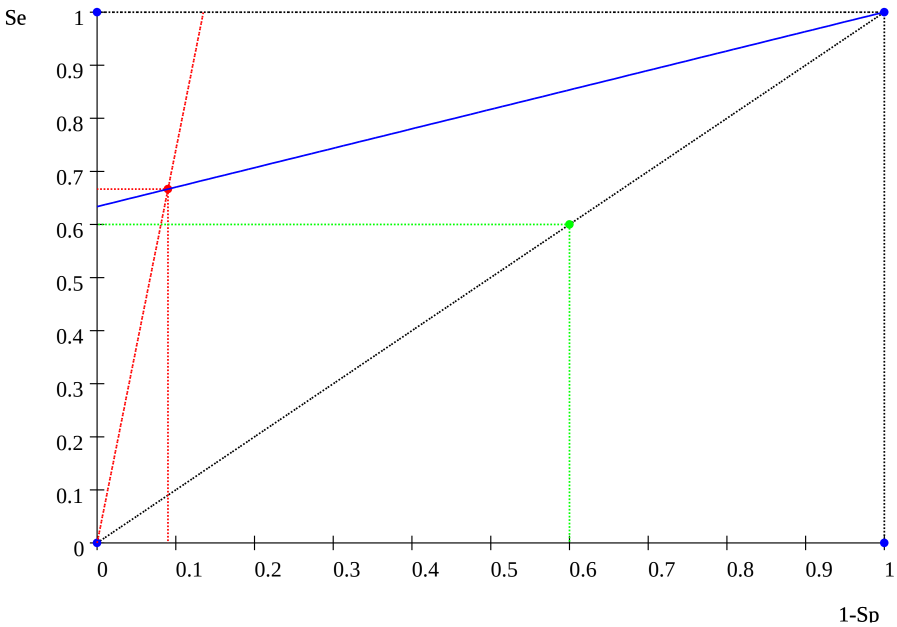


Figure 14 : ROC space

	D_+	D_-				
T_+	$\underline{Q}_+$	$(Q-\underline{Q})_+$	Q_+	$\text{PPV}=\dfrac{\underline{Q}_+}{Q_+}$	$\text{FDR}=\dfrac{(Q-\underline{Q})_+}{Q_+}$	$\text{LR}_+=\dfrac{\text{PPV}}{\text{FDR}}=\dfrac{\underline{Q}_+}{(Q-\underline{Q})_+}$
T_-	$\underline{Q}_-$	$(Q-\underline{Q})_-$	Q_-	$\text{FOR}=\dfrac{\underline{Q}_-}{Q_-}$	$\text{NPV}=\dfrac{(Q-\underline{Q})_-}{Q_-}$	$\text{LR}_-=\dfrac{\text{FOR}}{\text{NPV}}=\dfrac{\underline{Q}_-}{(Q-\underline{Q})_-}$
	$\underline{Q}$	$Q-\underline{Q}$	Q	$p=\dfrac{\underline{Q}}{Q}$		$\text{DOR}=\dfrac{\text{LR}_+}{\text{LR}_-}$
	$\text{TPR}=\dfrac{\underline{Q}_+}{\underline{Q}}$	$\text{FPR}=\dfrac{(Q-\underline{Q})_+}{Q-\underline{Q}}$	$s=\dfrac{Q_+}{Q}$			
	$\text{FNR}=\dfrac{\underline{Q}_-}{\underline{Q}}$	$\text{TNR}=\dfrac{(Q-\underline{Q})_-}{Q-\underline{Q}}$				

Table 22 : Calibration quality - main indicators

$$\text{ACC}:\ \text{ACC}=\frac{n_{++}+n_{--}}{n}=\frac{\underline{Q}^{+}+(Q-\underline{Q})^{-}}{Q}$$

$$\text{Adjusted ACC}\ \ \text{ACC}^{*}=\text{ACC}-\text{ACC}_{\text{TBSE}}=\frac{\underline{Q}^{+}}{Q}-\frac{(Q-\underline{Q})^{+}}{Q}$$

$$\text{Youden' J}:\ J=\frac{\underline{Q}^{+}}{\underline{Q}}-\frac{(Q-\underline{Q})^{+}}{Q-\underline{Q}}$$

$$\text{Kappa Score}:\ I_\kappa=\frac{\text{ACC}-p^{*}}{1-p^{*}}\ \text{ avec }\ p^{*}=ps_{+}+(1-p)(1-s_{+})$$

$$\text{Cohen Ratio}:\ \frac{I_\kappa}{I_\kappa^{\max}}=\frac{p_0-p^{*}}{p_0^{\max}-p^{*}}\ \text{ avec }\ p_0^{\max}=1+\max[p-s,s-p]$$

$$\text{Jaccard Index}\ \ \text{JAC}=\frac{|\underline{Q}\cap S_{+}|}{|\underline{Q}\cup S_{+}|}=\frac{\underline{Q}_{+}}{\underline{Q}_{+}+\underline{Q}_{-}+(Q-\underline{Q})_{+}}$$

8.3.3 More details

As with Affordability and the Incentive Effect, the tool provides more disaggregated information by producing confusion matrices:

- for group G1 only ("Households are not connected to the collective sanitation network and face the EP tariff only") and group G2 only ("Households are connected to the collective sanitation network and face the EPA tariff");
- for the group of most deprived Households (with a standard of living strictly below the poverty threshold entered by the user) and those who are not (with a standard of living at least equal to the threshold level entered by the user)

and :

- calculation of all the indicators listed above for each of these sub-groups;
- a basic breakdown, by sub-group, of some of these indicators

(with calculation of their contribution to the value taken by the indicator on the whole population) with for instance :

$$\mathrm{TPR}=\frac{\underline{Q}_+}{\underline{Q}}=\frac{\underline{Q}_1}{\underline{Q}_1+\underline{Q}_2}\times\frac{\underline{Q}_1^+}{\underline{Q}_1}+\frac{\underline{Q}_2}{\underline{Q}_1+\underline{Q}_2}\times\frac{\underline{Q}_2^+}{\underline{Q}_2}=\frac{\underline{Q}_1}{\underline{Q}}\times\mathrm{TPR}_1+\frac{\underline{Q}_2}{\underline{Q}}\times\mathrm{TPR}_2$$

$$\mathrm{PPV}=\frac{\underline{Q}_+}{Q_+}=\frac{Q_+^1}{Q_+^1+Q_+^2}\times\frac{\underline{Q}_+^1}{Q_+^1}+\frac{Q_+^2}{Q_+^1+Q_+^2}\times\frac{\underline{Q}_+^2}{Q_+^2}=\frac{Q_+^1}{Q_+}\times\mathrm{PPV}_1+\frac{Q_+^2}{Q_+}\times\mathrm{PPV}_2$$

$$\bar{S}_p=\frac{\left(Q-\underline{Q}\right)_+}{Q-\underline{Q}}=\frac{Q_1-\underline{Q}_1}{Q-\underline{Q}}\times\frac{\left(Q-\underline{Q}\right)_+^1}{\left(Q-\underline{Q}\right)_1}+\frac{Q_2-\underline{Q}_2}{\left(Q-\underline{Q}\right)}\times\frac{\left(Q-\underline{Q}\right)_+^2}{\left(Q-\underline{Q}\right)_2}=\frac{Q_1-\underline{Q}_1}{Q-\underline{Q}}\times\bar{S}_p^1+\frac{Q_2-\underline{Q}_2}{\left(Q-\underline{Q}\right)}\times\bar{S}_p^2$$

…

$$\mathrm{ACC}^*=\frac{\underline{Q}_+-\left(Q-\underline{Q}\right)_+}{Q}=\frac{Q_1}{Q}\times\frac{\underline{Q}_+^1-\left(Q-\underline{Q}\right)_+^1}{Q_1}+\frac{Q_2}{Q}\times\frac{\underline{Q}_+^2-\left(Q-\underline{Q}\right)_+^2}{Q_2}=\frac{Q_1}{Q}\times\mathrm{ACC}_1^*+\frac{Q_2}{Q}\times\mathrm{ACC}_2^*$$

$$\mathrm{JAC}=\frac{\underline{Q}_+}{\underline{Q}_++\underline{Q}_-+\left(Q-\underline{Q}\right)_+}=\frac{\underline{Q}_+^1+\underline{Q}_-^1+\left(Q-\underline{Q}\right)_+^1}{\underline{Q}_++\underline{Q}_-+\left(Q-\underline{Q}\right)_+}\times\mathrm{JAC}_1+\frac{\underline{Q}_+^2+\underline{Q}_-^2+\left(Q-\underline{Q}\right)_+^2}{\underline{Q}_++\underline{Q}_-+\left(Q-\underline{Q}\right)_+}\times\mathrm{JAC}_2$$

for the G1 vs. G2 partition (mutatis mutandis Poor vs. Non-poor partition). Finally, these decompositions are also implemented for the typology obtained by crossing the two criteria EP Service only (G1) vs. EPA service (G2) and Poor vs. Non Poor (with the 4 sub-populations G1-Poor, G2-Poor, G1-Non_Poor and G2-Non_Poor) and supplemented with 3 contigency tables showing (i) the breakdown of basic units which are rightly subsidised, (ii) the breakdown of basic units which are wrongly taxed and (iii) the breakdown of non basic units which are wrongly subsidised.

IX – EVALUATION – EQUITY

The Invoices module collects the information on the transfers implemented by the IBT which is evaluated/tested by the user. On this basis, the tool addresses the question of the equity of the EP/EPA service charge from two points of view. The first is a classical one that refers to the characterisation of the cross-subsidy system (see notably Komives et al. [2007], Barde & Lehmann P. [2014] and Fuente et al. [2016]). The second aims to answer a simple question that can be formulated as follows: does the IBT (considered by the user) increase or reduce income inequalities, compared with the reference pricing system TBSE, through its effects on the Lorenz curve, in general, and the Gini index in particular? Once these elements have been set, the tool breaks down the impact of subsidies and "taxes" (in fact, contributions to service funding) on Gini index with decompositions by factors, by groups and by groups and factors.

9.1 Measuring cross-subsidies

9.1.1 Data

A) Starting point Initially, the information available in the Invoices module is as follows (considering the case of an IBT4, to illustrate):

i	c_{i0}	c_{i1}	c_{i2}	c_{i3}	c_{i4}	s_{iq}	t_{iq}	$t_{iq}+s_{iq}$	$\mathbf{I}_i$	$c_{i°}+t_{iq}+s_{iq}$	$\mathbf{J}_i$
1											
2											
⋮											
n											

Table 23 : Breakdown of household contributions to service funding - IBT4

with:

(1) $c_{i0} = F - \frac{CF}{n}$ the contribution to service funding of household *i* on the fixed part/access fee, counted negatively in the case of a subsidy ($F < \frac{CF}{n}$) and positively in the case of a "tax" ($F > \frac{CF}{n}$);

Remark The variable c_{i0} is "Group-specific", with a common value to all household subscribers to the "EPA" service (Group 2), with $c_{i0} = F_{\text{EP}} + F_{\text{A}} - \frac{CF_{\text{EP}}}{n} - \frac{CF_{\text{A}}}{n_{\text{A}}}$, and a common but different value to all subscribers to the "EP" service only (Group 1), with $c_{i0} = F_{\text{EP}} - \frac{CF_{\text{EP}}}{n}$.

(2) $c_{ij} = \left(\pi_j - c\right) \times q_i^j$ the contribution to service funding from household *i* in consumption block *j*, *j* varying (here) from 1 to 4, which is counted negatively in the case of a subsidy ($\pi_j < c$) and positively in the case of a "tax" ($\pi_j > c$);

(3) q_i^j the consumption of block *j* of household *i* with, for instance, $q_i^1 = k_1$, $q_i^2 = q_i - k_1$, $q_i^3 = q_i^4 = 0$ for a household *i* whose consumption q_i is located in block 2;

(4) s_{iq} the sum of gross subsidies, counted (here) negatively, received by household *i* on all its consumption q_i :

$$s_{iq} = \min\left[c_{i1},0\right] + \min\left[c_{i2},0\right] + \min\left[c_{i3},0\right] + \min\left[c_{i4},0\right] = \sum_{j=1}^{p=4} \min\left[c_{ij},0\right]$$

(5) t_{iq} the sum of gross “taxes”, counted positively, which are levied on all consumption q_i of household *i*, hence the margin generated on household consumption *i*, which then takes the form of a gross contribution to service funding:

$$s_{iq} = \max\left[c_{i1},0\right] + \max\left[c_{i2},0\right] + \max\left[c_{i3},0\right] + \max\left[c_{i4},0\right] = \sum_{j=1}^{p=4} \max\left[c_{ij},0\right]$$

(6) $t_{iq} + s_{iq}$ the net contribution to service funding of consumption q_i from household *i*, counted negatively in the case of a subsidy[33];

(7) $\mathbf{I}_i$ a dummy variable that takes the value 1 if the consumption of household *i* is a net contributor to service funding ($t_{iq} + s_{iq} > 0$) and 0 otherwise ($t_{iq} + s_{iq} \leq 0$);

(8) $c_{i0} + t_{iq} + s_{iq}$ the net contribution to service funding, counted negatively in the case of a subsidy, from the household *i* with the Subsidy/Taxation on the Access Fee Included[34];

(9) $\mathbf{J}_i$ a dummy variable that takes the value 1 if the household is a net contributor to service funding ($c_{i0} + t_{iq} + s_{iq} > 0$) and 0 otherwise ($c_{i0} + t_{iq} + s_{iq} \leq 0$).

B) Exclusion errors in values on basic consumption Feeding the algorithm (which leads to the production of Table 23) with basic consumption $\underline{q}_i$ of household *i* as an input enables to identify:

(1) the amounts of the subsidies (which are rightly paid)

(2) the amounts of taxes (which are wrongly levied)

on water basic consumptions with an output that is precisely as follows:

i	$\underline{c}_{i0}$	$\underline{c}_{i1}$	$\underline{c}_{i2}$	$\underline{c}_{i3}$	$\underline{c}_{i4}$	$s_{i\underline{q}}$	$t_{i\underline{q}}$	$t_{i\underline{q}} + s_{i\underline{q}}$	$\underline{\mathbf{I}}_i$	$\underline{c}_{i0} + t_{i\underline{q}} + s_{i\underline{q}}$	$\underline{\mathbf{J}}_i$
1											
2											
⋮											
n											

Table 24 : Identification of exclusion errors in values - IBT4

[33] A positive value for $t_{iq} + s_{iq}$ indicates that the household is a net contributor through its consumption to service funding, i.e. the level of its consumption means that it brings in "taxes" more than it costs in subsidies.

[34] A positive value for $c_{i0} + t_{iq} + s_{iq}$ indicates that the household is a net contributor to the financing of the service, i.e. the household brings more in taxes than it costs in subsidies Axxess Fee Included.

The definitions of the variables are identical to those given for the Table 23, page 101, except that they refer to basic consumption. As a result:

- the variable $s_{i\underline{q}}$ gives the total subsidies, counted negatively, which are allocated by the tariff on the basic consumption of household *i*, Access Fee Excluded,
- the variable $t_{i\underline{q}}$ gives the total taxes which are levied by the tariff on the basic consumption of household *i*, Access Fee Excluded.

This last variable represents **a gross exclusion error**, in value, borne by the household *i* on its basic consumption $\underline{q}_i$ with the implementation of the tariff (this error takes the value 0 when the entire basic consumption of household *i* is subsidised). The $t_{i\underline{q}} + s_{i\underline{q}}$ variable, when positive, then reflects **a net error** of the same nature, and the $c_{i0} + t_{\underline{q}} + s_{\underline{q}}$ variable, when positive, **a net** (exclusion) **error in value Access Fee Included** (meeting the household's needs requires access to the service, and therefore payment of the access fee).

C) Inclusion errors in values on non-basic consumption The calculation of differences between the items c_{ij} of Table 23 and the items $\underline{c}_{ij}$ of Table 24 allows to identify:

(*i*) the amounts of subsidies that are wrongly paid on non-basic consumption in block *j*, with the values. $c_{ij} - \underline{c}_{ij} < 0$;

(*ii*) the amounts of the "taxes" that are rightly levied on non-basic consumption in block *j* with the values $c_{ij} - \underline{c}_{ij} > 0$.

By way of illustration, we have in particular:

- For a household whose basic consumption $\underline{q}$ is in block 1 and consumption *q* is in block 1:

c_{i0}	c_{i1}	c_{i2}	c_{i3}	c_{i4}
$F - \frac{CF}{n}$	$(\pi_1 - c)q_i$	0	0	0
$\underline{c}_{i0}$	$\underline{c}_{i1}$	$\underline{c}_{i2}$	$\underline{c}_{i3}$	$\underline{c}_{i4}$
$F - \frac{CF}{n}$	$(\pi_1 - c)\underline{q}_i$	0	0	0
$c_{i0} - \underline{c}_{i0}$	$c_{i1} - \underline{c}_{i1}$	$c_{i2} - \underline{c}_{i2}$	$c_{i3} - \underline{c}_{i3}$	$c_{i4} - \underline{c}_{i4}$
0	$(\pi_1 - c)(q_i - \underline{q}_i)$	0	0	0

- For a household whose basic consumption $\underline{q}$ is in block 1 and consumption *q* is in block 2:

c_{i0}	c_{i1}	c_{i2}	c_{i3}	c_{i4}
$F - \frac{CF}{n}$	$(\pi_1 - c)k_1$	$(\pi_2 - c)(q_i - k_1)$	0	0
$\underline{c}_{i0}$	$\underline{c}_{i1}$	$\underline{c}_{i2}$	$\underline{c}_{i3}$	$\underline{c}_{i4}$
$F - \frac{CF}{n}$	$(\pi_1 - c)\underline{q}_i$	0	0	0
$c_{i0} - \underline{c}_{i0}$	$c_{i1} - \underline{c}_{i1}$	$c_{i2} - \underline{c}_{i2}$	$c_{i3} - \underline{c}_{i3}$	$c_{i4} - \underline{c}_{i4}$
0	$(\pi_1 - c)(k_1 - \underline{q}_i)$	$(\pi_2 - c)(q_i - k_1)$	0	0

- For a household whose basic consumption $\underline{q}$ is in block 2 and consumption q is in block 3:

c_{i0}	c_{i1}	c_{i2}	c_{i3}	c_{i4}
$F-\frac{CF}{n}$	$(\pi_1-c)k_1$	$(\pi_2-c)(k_2-k_1)$	$(\pi_3-c)(q_i-k_2)$	0
$\underline{c}_{i0}$	$\underline{c}_{i1}$	$\underline{c}_{i2}$	$\underline{c}_{i3}$	$\underline{c}_{i4}$
$F-\frac{CF}{n}$	$(\pi_1-c)k_1$	$(\pi_2-c)(\underline{q}_i-k_1)$	0	0
$c_{i0}-\underline{c}_{i0}$	$c_{i1}-\underline{c}_{i1}$	$c_{i2}-\underline{c}_{i2}$	$c_{i3}-\underline{c}_{i3}$	$c_{i4}-\underline{c}_{i4}$
0	0	$(\pi_2-c)(k_2-\underline{q}_i)$	$(\pi_3-c)(q_i-k_2)$	0

- …

with elements (when they are different from 0) that will be negative for the subsidised blocks ($\pi_j < c$) and positive for the blocks that are taxed/subject to a contribution to service funding ($\pi_j > c$). Figure 15 and Figure 16, on the following pages, show the last two cases for the Saint Paul EP tariff and an unit variable cost estimated at €1.50.

The result of this operation ("table subtraction") is to produce a Table 25 with:

i	$c_{i0}^{q_i-\underline{q}_i}$	$c_{i1}^{q_i-\underline{q}_i}$	$c_{i2}^{q_i-\underline{q}_i}$	$c_{i3}^{q_i-\underline{q}_i}$	$c_{i4}^{q_i-\underline{q}_i}$	$s_{i,q-\underline{q}}$	$t_{i,q-\underline{q}}$	$t_{i,q-\underline{q}}+s_{i,q-\underline{q}}$	$\mathbf{1}_i^{q-\underline{q}}$
1									
2									
⋮									
n									

Table 25 : Identification of inclusion errors in value - IBT4

The variables $c_{ij}^{q-\underline{q}} = c_{ij} - \underline{c}_{ij}$ when they are negative measure an inclusion error in value in block *j* and the variable $s_{i,q-\underline{q}}$ (which corresponds to the sum of the $c_{ij}^{q-\underline{q}} = c_{ij} - \underline{c}_{ij}$ when these values are negative) the amount of the subsidy, counted negatively, which is wrongly paid on the non-basic consumption of household *i*.

The quantity $t_{i,q-\underline{q}}$ (which corresponds to the sum of the $c_{ij}^{q-\underline{q}} = c_{ij} - \underline{c}_{ij}$ when these values are positive) gives the amount of 'taxes' that are rightly levied on the non-basic consumption of household *i* (which acts as a targeted tax base) and the sum $t_{i,q-\underline{q}} + s_{i,q-\underline{q}}$, the amount of the net inclusion error incurred by subsidising the non-basic consumption $q_i - \underline{q}_i$ of household *i* when it is negative, the amount of the net contribution to service funding that is generated on the non-basic consumption of household *i* when it is positive.

The dummy variable $\mathbf{1}_i^{q-\underline{q}}$, when it takes the value 1, indicates that the household does ultimately contribute to the financing of the service on its non-basic consumption.

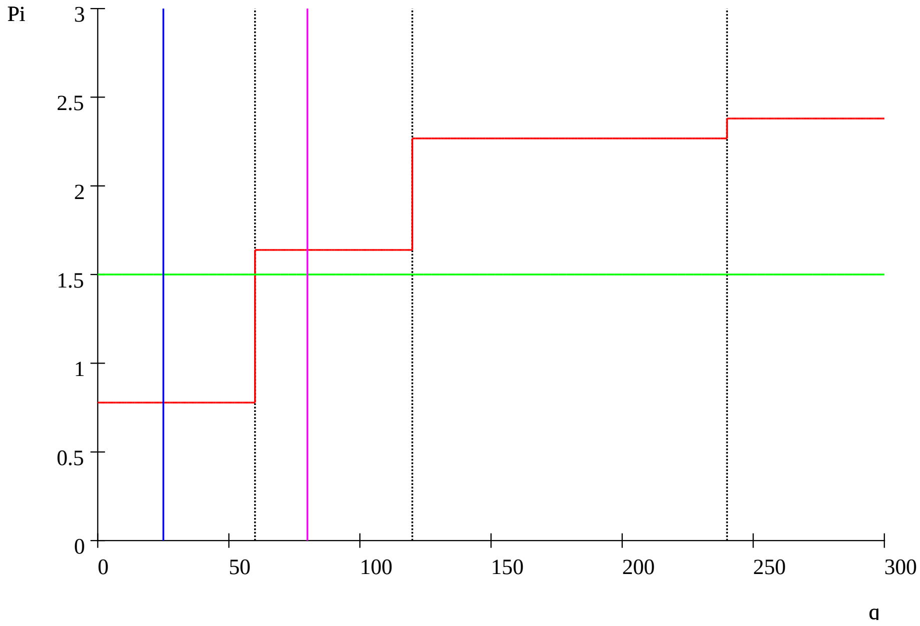


Figure 15 : Subsidies and "taxes" on consumption - EP *tariff for Saint Paul with* $\left(\underline{q}_i, q_i\right) = (25{,}80)$

- Consomption breakdown:

$$q_i = 80 = q_i^1 + q_i^2 + q_i^3 + q_i^4 = 60 + 20 + 0 + 0$$
$$\underline{q}_i = 25 = \underline{q}_i^1 + \underline{q}_i^2 + \underline{q}_i^3 + \underline{q}_i^4 = 25 + 0 + 0 + 0$$
$$q_i - \underline{q}_i = 80 - 25 = \left(q - \underline{q}\right)_i^1 + \left(q - \underline{q}\right)_i^2 + \left(q - \underline{q}\right)_i^3 + \left(q - \underset{\sim}{q}\right)_i^4 = 35 + 20 + 0 + 0 = 55$$

- Subsidies:

$$s_{q_i} = \left(\pi_1 - c\right) \times k_1 = \left(0.778 - 1.5\right) \times 60 = -43.32$$
$$s_{\underline{q}_i} = \left(\pi_1 - c\right) \times \underline{q}_i = \left(0.778 - 1.5\right) \times 25 = -18.05$$
$$s_{\left(q-\underline{q}\right)_i} = \left(\pi_1 - c\right) \times \left(q - \underline{q}\right)_i^1 = \left(0.778 - 1.5\right) \times \left(60 - 25\right) = -25.27$$

- Taxes:

$$t_{q_i} = \left(\pi_2 - c\right) \times \left(q_i - k_1\right) = \left(1.639 - 1.5\right) \times \left(80 - 60\right) = 2.78$$
$$t_{\underline{q}_i} = 0$$
$$t_{\left(q-\underline{q}\right)_i} = \left(\pi_2 - c\right) \times \left(q - \underline{q}\right)_i^2 = \left(1.639 - 1.5\right) \times 20 = 2.78$$

- Net positions:

$$t_{q_i} + s_{q_i} = 2.78 + \left(-43.32\right) = -40.54$$
$$t_{\underline{q}_i} + s_{\underline{q}_i} = 0 + \left(-18.05\right) = -18.05$$
$$t_{\left(q-\underline{q}\right)_i} + s_{\left(q-\underline{q}\right)_i} = 2.78 + \left(-25.27\right) = -22.49$$

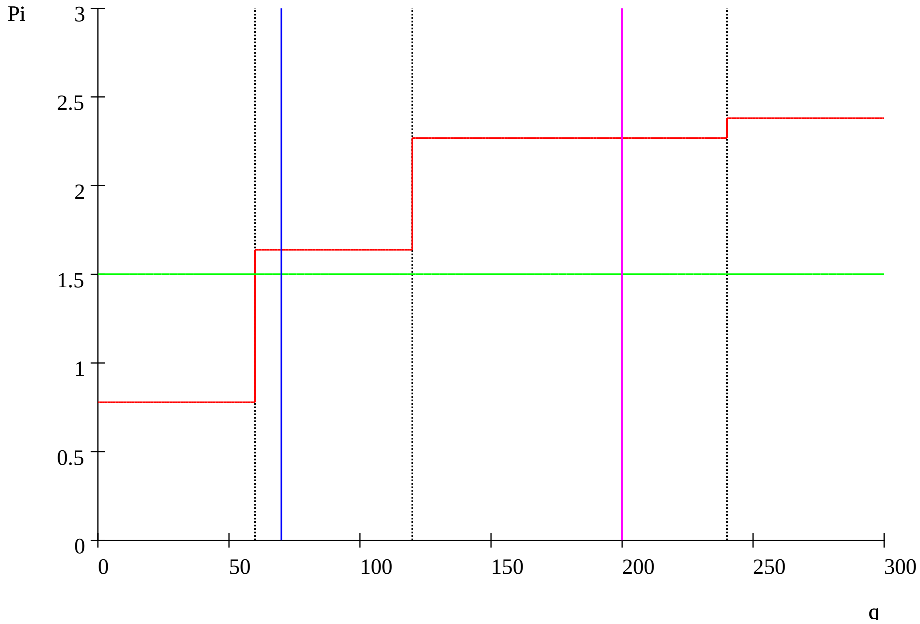


Figure 16 : Subsidies and "taxes" on consumption - EP *tariff for Saint Paul with* $\left(\underline{q}_i, q_i\right) = (70, 200)$

- Consomption breakdown:

$$q_i = 200 = q_i^1 + q_i^2 + q_i^3 + q_i^4 = 60 + 60 + 80 + 0$$
$$\underline{q}_i = 70 = \underline{q}_i^1 + \underline{q}_i^2 + \underline{q}_i^3 + \underline{q}_i^4 = 60 + 10 + 0 + 0$$
$$q_i - \underline{q}_i = 200 - 70 = \left(q_i^1 - \underline{q}_i^1\right) + \left(q_i^2 - \underline{q}_i^2\right) + \left(q_i^3 - \underline{q}_i^3\right) + \left(q_i^4 - \underline{q}_i^4\right) = 0 + 50 + 80 + 0 = 130$$

- Subsidies:

$$s_{q_i} = \left(\pi_1 - c\right) \times k_1 = \left(0.778 - 1.5\right) \times 60 = -43.32$$
$$s_{\underline{q}_i} = \left(\pi_1 - c\right) \times \underline{q}_i^1 = \left(0.778 - 1.5\right) \times 60 = -43.32$$
$$s_{\left(q-\underline{q}\right)_i} = \left(\pi_1 - c\right) \times \left(q - \underline{q}\right)_i^1 = \left(0.778 - 1.5\right) \times 0 = 0$$

- Taxes:

$$t_{q_i} = \left(\pi_2 - c\right) \times q_i^2 + \left(\pi_3 - c\right) \times q_i^3 = \left(1.639 - 1.5\right) \times 60 + \left(2.268 - 1.5\right) \times 80 = 69.78$$
$$t_{\underline{q}_i} = \left(\pi_2 - c\right) \times \underline{q}_i^2 = \left(1.639 - 1.5\right) \times 10 = 1.39$$
$$t_{\left(q-\underline{q}\right)_i} = \left(\pi_2 - c\right) \times \left(q - \underline{q}\right)_i^2 + \left(\pi_3 - c\right) \times \left(q - \underline{q}\right)_i^3 = \left(1.639 - 1.5\right) \times 50 + \left(2.268 - 1.5\right) \times 80 = 68.39$$

- Net positions :

$$t_{q_i} + s_{q_i} = 69.78 + \left(-43.32\right) = 26.46$$
$$t_{\underline{q}_i} + s_{\underline{q}_i} = 1.39 + \left(-43.32\right) = -41.93$$
$$t_{\left(q-\underline{q}\right)_i} + s_{\left(q-\underline{q}\right)_i} = 68.39 + \left(-0\right) = 68.39$$

9.1.2 Net subsidies and taxes

9.1.2.1 Descriptive Statistics

The tool starts by giving the main descriptive statistics on:

(*i*) the subsidy on the acces fee and (ii) the (possible but unlikely) taxation on the access fee:

$$\left(c_0\right)_- = \begin{cases} \frac{CF}{n} - F & if \quad F < \frac{CF}{n} \\ 0 & if \quad F \geq \frac{CF}{n} \end{cases} \tag{9.1}$$

$$\left(c_0\right)_+ = \begin{cases} F - \frac{CF}{n} & if \quad F > \frac{CF}{n} \\ 0 & if \quad F \leq \frac{CF}{n} \end{cases} \tag{9.2}$$

(iii) the net subsidies granted on household water consumption:

$$\left(t_q + s_q\right)_- = \begin{cases} -\left(t_q + s_q\right) & if \quad t_q + s_q < 0 \\ 0 & if \quad t_q + s_q \geq 0 \end{cases} \tag{9.3}$$

(iv) the net contributions to service funding from household water consumption:

$$\left(t_q + s_q\right)_+ = \begin{cases} t_q + s_q & if \quad t_q + s_q > 0 \\ 0 & if \quad t_q + s_q \leq 0 \end{cases} \tag{9.4}$$

(v) the net subsidy including access fee :

$$\left(c_{i0} + t_q + s_q\right)_- = \begin{cases} -\left(c_{i0} + t_q + s_q\right) & if \quad c_{i0} + t_q + s_q < 0 \\ 0 & if \quad c_{i0} + t_q + s_q \geq 0 \end{cases} \tag{9.5}$$

(vi) the net contribution to the financing of the service including access fee:

$$\left(c_{i0} + t_q + s_q\right)_- = \begin{cases} -\left(c_{i0} + t_q + s_q\right) & if \quad c_{i0} + t_q + s_q < 0 \\ 0 & if \quad c_{i0} + t_q + s_q \geq 0 \end{cases} \tag{9.6}$$

for the general service EP / EPA. This information is provided for the household population as a whole (with the truncated variables given above), and for the sub-population of household for which no 0 is recorded (actual values). In this case, the table also shows the weight (%) of the sub-population concerned. See Table 26, page 108, for the net subsidies, and Table 27, page 109, for the net contributions to service funding, for a numerical illustration.

Table 26 : Net subsidies - Descriptive statistics

	Total Population			Beneficiary Population		
	Access Fee	"DAE"	"DAI"	Access Fee	DAE	DAI
%	100.00	6.33	62.23	**	**	**
Mean	49.98	0.21	16.77	49.98	3.30	26.95
Median	28.93	0.00	8.70	28.93	2.56	23.47
Min	28.93	0.00	0.00	28.93	0.15	0.10
Max	74.85	7.98	81.36	74.85	7.98	81.36
Q1	28.93	0.00	0.00	28.93	1.50	10.80
Q3	74.85	0.00	29.03	74.85	5.59	36.86
D1	28.93	0.00	0.00	28.93	0.25	4.38
D9	74.85	0.00	50.25	74.85	6.84	61.41
F (Mean)	54.1	94.2	63.7	54.1	61.7	54.0
Variance	523.4846	1.0436	428.9656	523.4846	6.2781	414.9678
Standard deviation	22.88	1.02	20.71	22.88	2.51	20.37
MAPE	22.80	0.39	16.91	22.80	2.22	16.60
Coeff of Variation	0.458	4.888	1.235	0.458	0.759	0.756
Interquartile range	45.92	0.00	29.03	45.92	4.10	26.06
Interdecile range	45.92	0.00	50.25	45.92	6.59	57.03
Yule coefficient	1.00	n.a.	0.40	1.00	0.48	0.03
Gini index	22.81	96.4	63.8	22.8	42.6	41.9
Schutz coefficient	22.8	93.8	50.4	22.8	33.6	30.8
Interdecile ratio	2.59	n.a.	n.a.	2.59	27.24	14.03

Table 27 : Net taxes (margins) - Descriptive statistics

	Totak Population			Beneficiary Population		
	Access Fee	"DAE"	"DAI"	Access Fee	"DAE"	"DAI"
%	0.00	93.67	37.77	**	**	**
Mean	0.00	49.87	16.45	0.00	53.24	43.54
Median	0.00	39.54	0.00	0.00	42.17	27.68
Min	0.00	0.00	0.00	0.00	0.05	0.75
Max	0.00	321.45	246.60	0.00	321.45	246.60
Q1	0.00	15.93	0.00	0.00	19.36	13.33
Q3	0.00	68.96	18.73	0.00	69.97	62.56
D1	0.00	1.98	0.00	0.00	7.65	4.43
D9	0.00	108.14	57.25	0.00	113.17	101.23
F (Mean)	0	59.2	73.3	0	59.6	64.6
Variance	0.0000	2276.6396	1121.8347	0.0000	2251.0719	1790.3405
Standard deviation	0.00	47.71	33.49	0.00	47.45	42.31
MAPE	0.00	34.72	22.40	0.00	34.26	31.95
Coeff of Variation	n.a.	0.957	2.037	n.a.	0.891	0.972
Interquartile range	0.00	53.03	18.73	0.00	50.61	49.23
Interdecile range	0.00	106.15	57.25	0.00	105.52	96.80
Yule coefficient	n.a.	0.11	1.00	n.a.	0.10	0.42
Gini index	n.a.	48.4	80.9	n.a.	45.0	49.4
Schutz coefficient	n.a.	34.8	68.1	n.a.	32.2	36.7
Interdecile ratio	n.a.	54.56	n.a.	n.a.	14.79	22.86

9.1.2.2 Relative benefit and contribution curves

The tool next uses a diagram often used in the literature (see notably Estupiñán et al. [2007]) to assess the redistributive/anti-redistributive nature of subsidies and "taxes" with:

(i) the relative benefit distribution curves for, successively, the net subsidy "Access Fee Excluded" and the net subsidy "Access Fee Included";

(ii) the relative contribution distribution curves for, successively, the net contributions to service funding "Access Fee Excluded" (these are the margins generated on consumption) and the net contributions to service funding "Access Fee Included" (these are the margins generated on household).

These curves show the concentration of the variable of interest in relation to the standard of living of household (domestic subscribers). See Figure 17 on page 111 for an illustration and Appendix 5, which gives details of the construction of these diagrams focussing on the net subsidy "Acces Fee Excluded" case.

These infographics are next completed by the calculation of two scalar indices (per variable of interest) with:

(1) the calculation of quasi-Gini index which can take on negative values (see Figure 17.1 and Figure 17.2) with a system (of cross-subsidies) which is then described as redistributive,

(2) the calculation of Omega ratio, defined as the ratio between the share of subsidies (*mutatis mutandis* of taxes) that goes to poor household divided by the poverty rate, here the percentage of poor households,

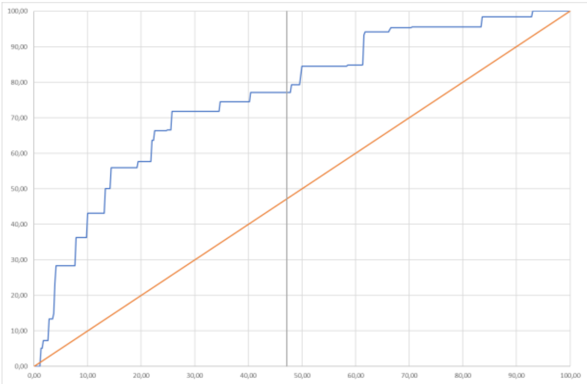
(both indicators are also computed for the most likely case in which the tariff would subsidise the right of access, as displayed in the table below). The first indicator makes it possible to assess the overall redistributive/anti-redistributive nature of (net) subsidies and (net) contributions to service funding (in particular, the closer the indicator is to -1, the higher the proportion of (in this case net) subsidies going to the poorest households; conversely, the closer the indicator is to 1, the higher the proportion of (in this case net) subsidies going to the wealthiest households). The second indicator, which also corresponds to the ratio of the average subsidies received by poor households to the average subsidies received by the population as a whole, measures the social fairness of the tariff mechanism (in relation to the pivot value $\Omega = 1$ which describes a situation in which subsidies and taxes are randomly distributed within the population of households).

This information is shown in the tool for the population of the household (see table below) and, also, for the population of individuals with the calculation of per capita values.

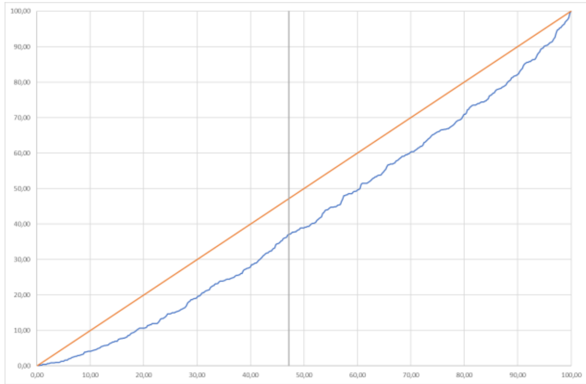
	Sub DA	Sub DAE	Sub DAI	Margin_DAE	Margin_DAI
Mean	49.98	0.21	16.77	49.87	16.45
Mean Poor	51.25	0.34	21.46	39.26	9.12
Ω ratio	1.03	1.64	1.28	0.79	0.55
Quasi Gini Index	-1.09	-52.3	-23.5	17.4	35.0

Table 28 : Quasi-Gini Index and Omega ratios

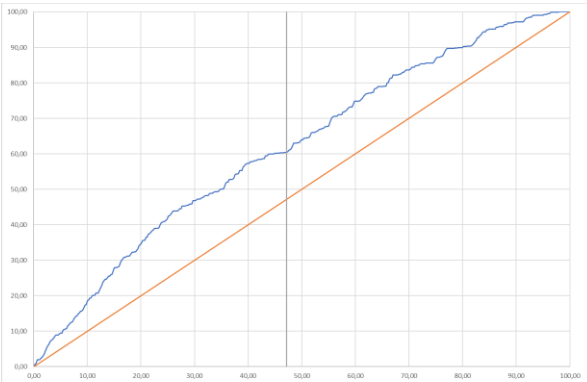
Figure 17 : Relative Benefit and Contribution Distribution Curves (Ω ratio values and Quasi-Gini Index (visualization))



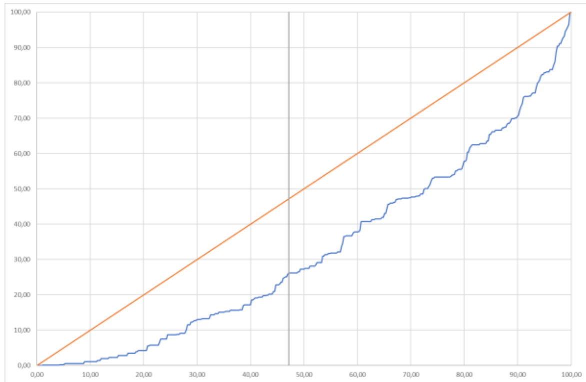
17.1: showing the relative beneficiary curve for DAE subsidies



17.3: showing the relative contribution curve for DAE taxations



17.2: showing the relative beneficiary curve for DAI subsidies



17.4: Showing the relative contribution curve for DAI taxations

9.1.2.3 Absolute beneficiary and contributor curves

The tool concludes this initial analysis on the distribution of subsidies and net "taxes" along the spectrum of living standards with the production of a second infographic frequently used in the literature (see notably Gómez-Lobo & Contreras [2003]). Originally, this diagram is used to measure and visualise the exclusion and inclusion errors for households relative to a target population defined here as the population of poor households. The exclusion error is then defined as the percentage of the population of poor households who do not receive any subsidy, and the inclusion error as the percentage of the population of beneficiaries who should not receive a subsidy, i.e. who do not belong to the target population. As apparent, the implementation of this approach focuses attention on one of the dimensions of the social component of the tariff, namely its ability to reach poor households through the cross-subsidy system.

To calculate these household inclusion and exclusion errors, one needs to identify poor families and those who are not in the Subscribers file of the Population module. This operation is carried out on the basis of the value set by the user for the poverty threshold (in the Social Data section of the General Data tab), with the creation of a dumy variable $\mathbf{1}_i^{\text{Poor}}$ that takes the value 1 if the household is in poverty and 0 otherwise. On this basis, the tool constructs the Absolute Beneficiary Curve which, in the case of net subsidies "Access Fee Excluded" $\left(t_q + s_q\right)_-$, (mutatis mutandis the net subsidies "Access Fee Included" $\left(c_0 + t_q + s_q\right)_-$) graphs the cumulative number of beneficiaries against the normalised rank of households in Pen's parade of living standards with:

$$B_h = \frac{1}{n}\sum_{i^*=1}^{h}\mathbf{1}_{i^*} \tag{9.7}$$

for $h = 1,\ldots,n$ with $\mathbf{1}_{i^*}$ a dummy variable that takes the value 1 if the household i^* is a net beneficiary of the Acces Fee Excluded subsidy system and 0 otherwise (as with the relative benefit distribution curve, households are ranked in ascending order of their standard of living). The graph of this function is shown in blue in Figure 18[35], on next page. Finalising the diagram requires:

- to display the 45° line and the vertical line of poverty $F = F_{\text{Poor}}$ (in green in Figure 18);
- to draw, at the point of intersection (marked A) between the 45° line and the poverty line $F = F_{\text{Poor}}$, the horizontal segment AB with $\text{A} = (F_{\text{Poor}}, F_{\text{Poor}})$ and $\text{B} = (1, F_{\text{Poor}})$.

The curve formed by the broken line OAB (in red in Figure 18) constitutes the Perfect Targeting Curve, in which only $100F_{\text{Poor}}\%$ of poor households benefit from subsidies. The latter enables to vizualise:

(1) the exclusion error with the ratio:

[35] The variable *r* (which is similar to an increasing cumulative frequency) corresponds to the standardised rank of Households in the Pen's parade of living standards.

$$\frac{A-C}{A}=\frac{F_{Poor}-B(F_{Poor})}{F_{Poor}} \quad (9.8)$$

with $B(F_{Poor})$ the percentage of poor households that actually benefit from the subsidy system (in the present case, net subsidy Access Fee Excluded);

(2) the inclusion error with the ratio:

$$\frac{D-C}{1}=\frac{B(1)-B(F_{Poor})}{B(1)} \quad (9.9)$$

where $B(1)$ is the percentage of households that actually receive a subsidy.

One of the advantages of this infographic is that it makes it possible to determine the size of the inclusion and exclusion errors for households at any value of the poverty line and, in particular, for households close to the poverty line (set by the user), and, therefore, in a vulnerable situation.

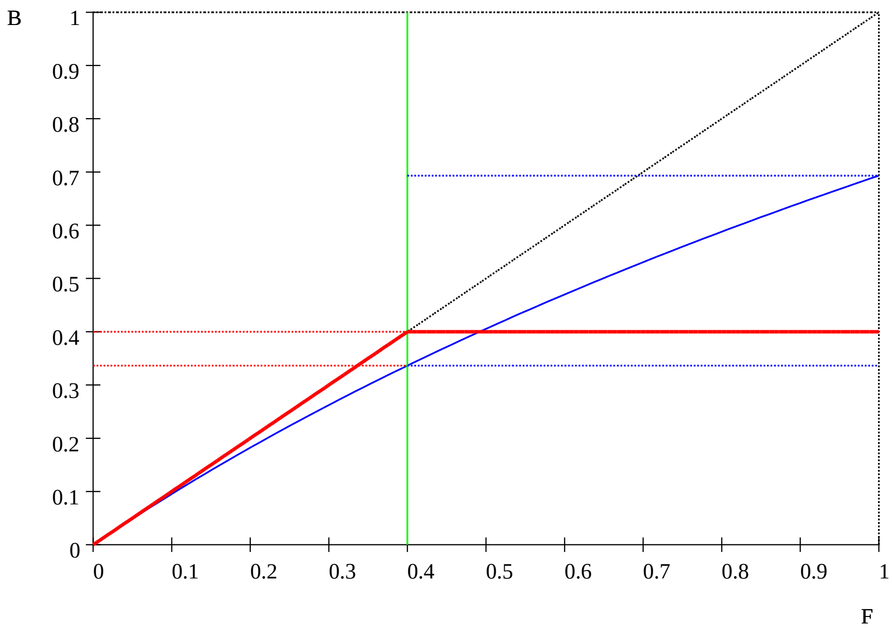
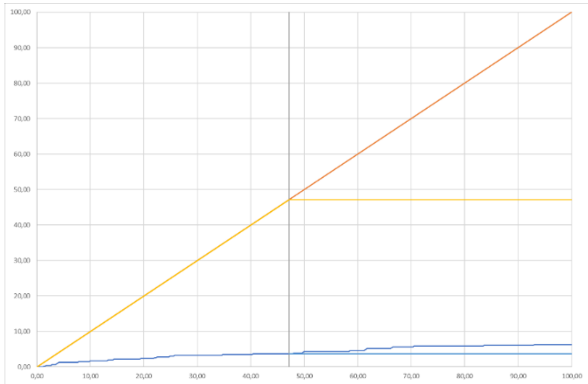


Figure 18 : Cumulative Beneficiary Curve for DAE net subsidies

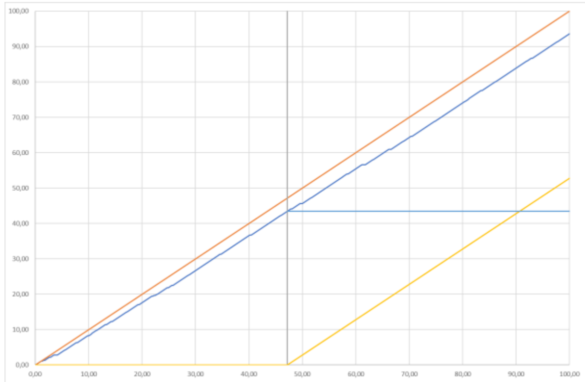
This inforgraphic is then produced for, successively, (i) the net subsidies Access Fee Excluded, (ii) the net subsidies Access Fee Included and, also, (iii) the margins generated on household consumption (net contributions to service funding Access Fee Excluded) and (iv) the margins generated on Households (net contributions to service funding Access Fee Included). which are generated by the IBT that is tested/assessed by the user[36]. See the set Figure 19 on next page.

[36] For these two variables, the Perfect Targeting Curve changes: it merges with the x-axis to the left of the Poverty line, and increases linearly, at a rate of 1 to 1, with the normalised rank of households to the right of the Poverty line. For the main, these curves simply state that poor households should not be charged on their consumption / on their basic consumption and the Access Fee, and therefore lose out compared to the TBSE (for the given IBT consumption).

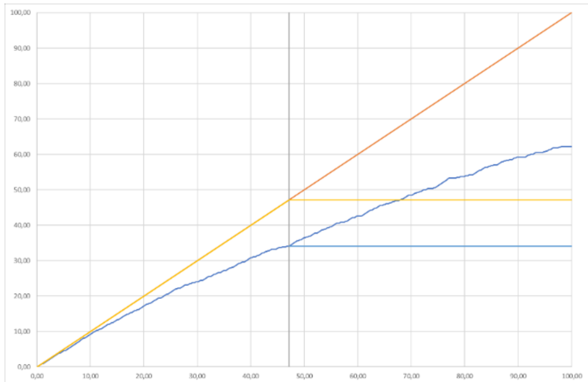
Figure 19 : Absolute Benefit and Contributor Curves



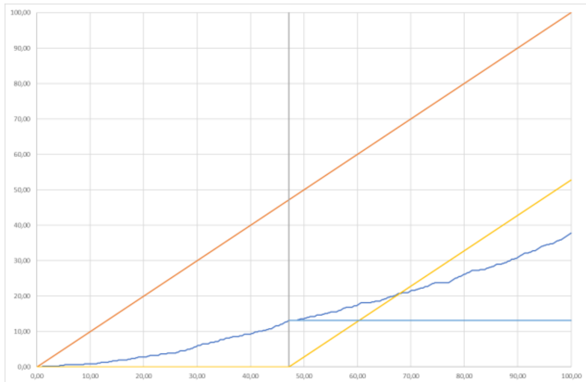
19.1: showing the absolute beneficiary curve for DAE subsidies



19.3: showing the absolute contributor curve for DAE taxations



19.2: showing the absolute beneficiary curve for DAE subsidies



19.4: Showing the absolute contributor curve for DAI taxations

The measurement of household exclusion and inclusion errors is next supplemented by the calculation of the leakage rate, which gives the percentage of the mass of net subsidy that does not go to the target population, in this case households living in poverty, and the "fiscal" burden for this category of the population (with the proportion of the mass of the net contributions to service funding, Access Fee Included and Access Fee Excluded, that are borne by poor households). This information is shown in Table 29 below. As apparent, insofar as one deals with net subsidies and net "taxations", Access Fee Included and Access Fee Excluded, the exclusion errors on the "DAE" and "DAI" subsidies correspond to the inclusion errors on the contributions to the "DAE" and "DAI" funding (modulo the case of poor households for whom the system would be neutral).

En %	Sub DA	Sub DAE	Sub DAI	Margin_DAE	Margin_DAI
EE	0	92.1	27.8	5.0	53.3
EI	52.8	5.0	53.3	92.1	27.8
Leakage rates	51.6	22.8	39.7		
Burden share				37.1	26.17

Table 29 : Inclusion/exclusion errors, leakage rates and burden share

To conclude, it should be noted that the tool offers the user the possibility of reproducing the entire analysis (Descriptive Statistics + Relative Benefit and Contribution Curves + Absolute Beneficiary and Contributor Curves) just presented in the case of the general "EP/EPA" service for, successively, the drinking water service, the collective sanitation service and the complete drinking water and collective sanitation service (which amounts to focusing on Group 2). For the sake of completeness, the user can also focus on households in groups 1 and 2 for the drinking water service alone.

9.1.3 Net errors of exclusion and inclusion in value

The tool reproduces all the elements in 9.1.2 for net subsidies and net taxes on basic and non-basic consumption (information on gross subsidies and gross taxes appears in the Cost Recovery field). See, by way of example, Table 31, Table 32, Figure 20 and Figure 21, pages 116 to 119 that relate to basic service as well as the table below which provides information on the inclusion and exclusion errors for basic service generated by the same tariff as the one evaluated with Table 29. The point is that a rather different (but complementary) reading is provided with, in particular, a tariff that now appears to be anti-redistributive. Finally, and as before, the user can focus on a service (EP, A, EPA) and, for sake of completeness, on households in group 1 and group 2 for the drinking water service (the examination of A and EPA services concerns group 2 only).

En %	Sub DA	Sub DAE	Sub DAI	Margin_DAE	Margin_DAI
Exclusion Err.	0	37.0	0	78.8	**
Inclussion Err.	100	79.8	100	37.0	**
Leakage rates	51.6	57.9	53.2		
Burden share				71.5	**

Table 30 : Inclusion/exclusion errors, leakage rates and burden share – Basic consumption

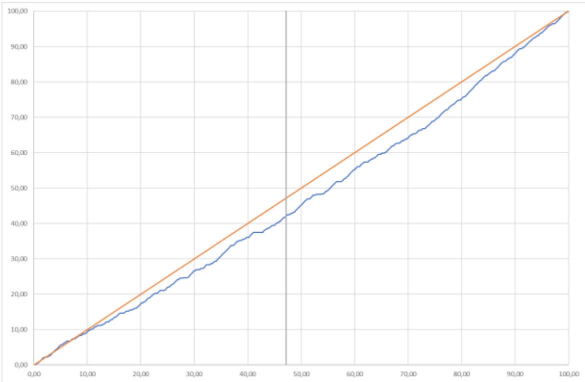
Table 31 : Net subsidies Basic Service - Descriptive statistics

	Totak Population			Beneficiary Population		
	Access Fee	"DAE"	"DAI"	Access fee	"DAE"	"DAI"
%	100.00	71.83	100.00	**	**	**
Mean	49.98	4.80	52.63	49.98	6.68	52.63
Median	28.93	7.10	36.82	28.93	7.50	36.82
Min	28.93	0.00	30.51	28.93	1.58	30.51
Max	74.85	8.22	83.07	74.85	8.22	83.07
Q1	28.93	0.00	36.52	28.93	7.00	36.52
Q3	74.85	7.69	74.05	74.85	7.69	74.05
D1	28.93	0.00	36.08	28.93	3.09	36.08
D9	74.85	7.80	79.49	74.85	7.88	79.49
F (Mean)	54.1	39.7	55.5	54.1	22.4	55.5
Variance	523.4846	11.4040	355.4662	523.4846	3.3242	355.4662
Standard deviation	22.88	3.38	18.85	22.88	1.82	18.85
MAPE	22.80	3.14	18.03	22.80	1.37	18.03
Coeff of Variation	0.458	0.704	0.358	0.458	0.273	0.358
Interquartile range	45.92	7.69	37.53	45.92	0.68	37.53
Interdecile range	45.92	7.80	43.41	45.92	4.79	43.41
Yule coefficient	1.00	-0.85	0.98	1.00	-0.46	0.98
Gini index	22.81	37.1	19.0	22.8	0.1	3.2
Schutz coefficient	22.8	32.7	17.1	22.8	10.3	17.1
Interdecile ratio	2.59	**	2.20	2.59	2.55	2.20

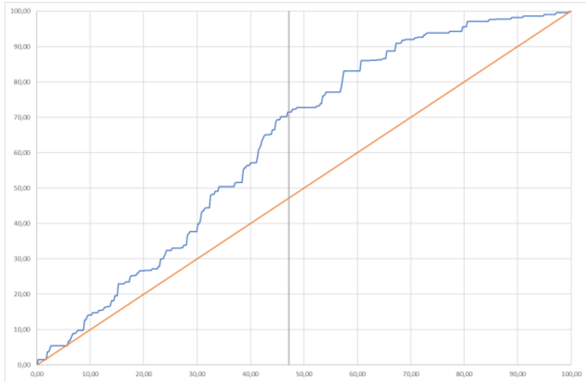
Table 32 : Net Tax (Margin) Basic Service - Descriptive Statistics

	Total Population			Beneficiary Population		
	Access Fee	"DAE"	"DAI"	Access fee	"DAE"	"DAI"
%	0.00	28.17	0.00	**	**	**
Mean	0.00	2.15	0.00		7.63	
Median	0.00	0.00	0.00		6.44	
Min	0.00	0.00	0.00		0.11	
Max	0.00	33.20	0.00		33.20	
Q1	0.00	0.00	0.00		0.80	
Q3	0.00	0.80	0.00		10.51	
D1	0.00	0.00	0.00		0.80	
D9	0.00	6.72	0.00		21.41	
F (Mean)	0	80.3	0		69.9	
Variance	0.0000	29.5759	0.0000		63.1990	
Standard deviation	0.00	5.44	0.00		7.95	
MAPE	0.00	3.33	0.00		5.90	
Coeff of Variation	n.a.	2.531	**		1.042	
Interquartile range	0.00	0.80	0.00		9.71	
Interdecile range	0.00	6.72	0.00		20.61	
Yule coefficient		1.00	**		-0.16	
			*			
Gini index		86.8	**		53.0	
Schutz coefficient		77.5	**		38.7	
Interdecile ratio		n.a.	**		26.76	

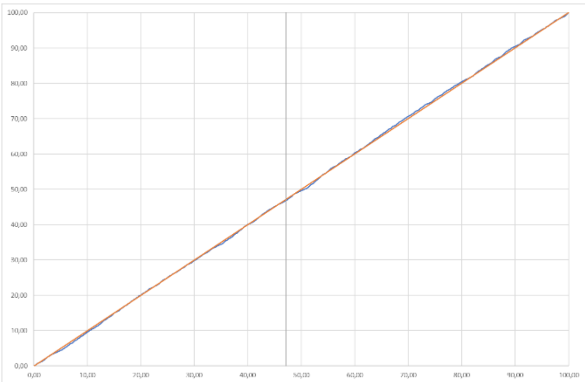
Figure 20 : Relative Benefit and Contribution Distribution Curves –Basic Service



20.1: showing the relative beneficiary curve for DAE subsidies



20.3: showing the relative contribution curve for DAE taxations



7

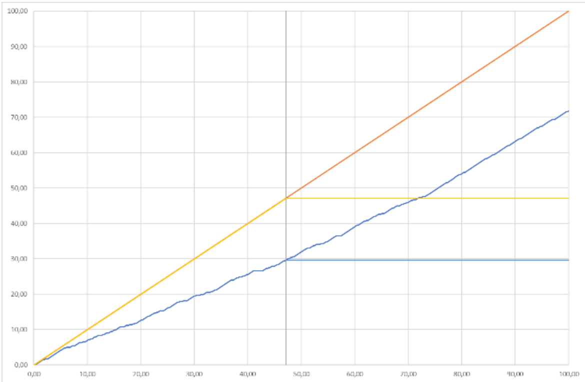
20.2: showing the relative beneficiary curve for DAI subsidies

No household is a net contributor to service funding

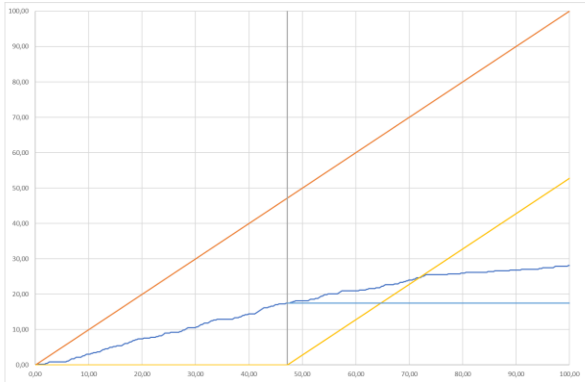
on its basic consumption.

18.4: Showing the relative contribution curve for DAI taxations

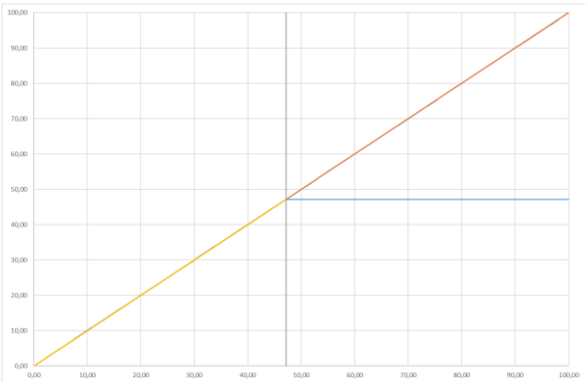
Figure 21 : Absolute Benefit and Contributor Curves – Basic Service



21.1: showing the absolute beneficiary curve for DAE subsidies



21.3: showing the absolute contributor curve for DAE taxations



21.2: showing the absolute beneficiary curve for DAI subsisides

No household is contributor to service funding

on its basic consumption Access Dee Included

21.4: Showing the absolute contributor curve for DAI taxations

9.2 The impact on inequality - Lorenz curves and Gini index

The main focus here is on the impact of progressive pricing on household income inequalities (Domestic Subscriber Approach). To this end, the tool uses the following methodology:

Stage 1 The first step is to compare (*i*) the Lorenz curve of household income that emerges with the IBT considered by the user and the IBT consumptions, for instance with an IBT4:

$$q_i^{\mathrm{IBT}} = q_i^d \left(F, \pi_1, \pi_2, \pi_3, \pi_4, k_1, k_2, k_3 \right) \tag{9.10}$$

with (*ii)* the Lorenz curve that emerges with the TBSE in which:

$$q_i^{\mathrm{TBSE}} = q_i^d \left(\tfrac{CF}{n}, c \right) \tag{9.11}$$

Under the IBT pricing scheme, the available income net of the water bill of the household *i* is given by:

$$R_i^{\mathrm{net\_IBT}} = R_i - T_i^{\mathrm{IBT}} = R_i - \left(F - D_i + \pi_i q_i^{\mathrm{IBT}} \right) \tag{9.12}$$

with D_i the value of Nordin's D ("virtual cashback") and π_i the marginal price faced by household *i*. If the water utility now applies a TBSE, we have:

$$R_i^{\mathrm{net\_TBSE}} = R_i - T_i^{\mathrm{TBSE}} = R_i - \left(\frac{CF}{n} + c \times q_i^{\mathrm{TBSE}} \right) \tag{9.13}$$

After ranking **for each pricing scheme** households by income level, net of the water bill, from the lowest one to the highest one, the tool:

- represents the two Lorenz curves, L_{IBT} and L_{TBSE} , of the net household income $R_i^{\mathrm{net\_IBT}} = R_i - T_i^{\mathrm{IBT}}$ and $R_i^{\mathrm{net\_TBSE}} = R_i - T_i^{\mathrm{TBSE}}$;

and next computes:

- the Gini index for the TBSE pricing scheme with the TBSE consumptions q_i^{TBSE} and the TBSE water bills $T_i^{\mathrm{TBSE}} = \frac{CF}{n} + c q_i^{\mathrm{TBSE}}$;
- the Gini index for the IBT pricing scheme with the IBT consumptions q_i^{IBT} and the IBT water bills $T_i^{\mathrm{IBT}} = F - D_i + \pi_i q_i^{\mathrm{IBT}}$.

The gross impact of the IBT on inequality (in household net income) is then measured by the change in the Gini index:

$$\Delta i = i_{\mathrm{IBT}} - i_{\mathrm{TBSE}} \tag{9.14}$$

Geometrically, this variation corresponds to the difference in area between the Lorenz curve of net income L_{IBT} , calculated with IBT consumption/invoices, and the Lorenz curve of net income L_{TBSE} calculated with TBSE consumption/invoices (it should be noted that, , in practice, the two curves are often very close to each other with a very small variation in the Gini index)

.

Step 2 Compared with the TBSE, the change in net income of the household *i* generated by the introduction of the IBT is, except for the sign, given by the change in water bill with:

$$\begin{aligned}\Delta R_i^{net} &= R_i^{\text{net\_IBT}} - R_i^{\text{net\_TBSE}} \\ &= R_i - \left(F - D_i + \pi_i \times q_i^{\text{IBT}}\right) - \left[R_i - \left(\frac{CF}{n} + cq_i^{\text{TBSE}}\right)\right] \\ &= -\left[\left(F - D_i + \pi_i \times q_i^{\text{IBT}}\right) - \left(\frac{CF}{n} + c \times q_i^{\text{TBSE}}\right)\right]\end{aligned} \qquad (9.15)$$

This variation in income can be expressed as:

$$\Delta R_i^{net} = -\left(F - \frac{CF}{n} + (\pi_i - c) \times q_i^{\text{IBT}} - D_i\right) + c \times \left(q_i^{\text{TBSE}} - q_i^{\text{IBT}}\right) = A_i + B_i \qquad (9.16)$$

$$A_i \equiv -\left(F - \frac{CF}{n} + (\pi_i - c) \times q_i^{\text{IBT}} - D_i\right) = -\left(c_{i0} + t_{i,q} + s_{i,q}\right) \qquad (9.17)$$

$$B_i = c \times \left(q_i^{\text{TBSE}} - q_i^{\text{IBT}}\right) \qquad (9.18)$$

with $c_{i0} = F - \frac{CF}{n}$ the subsidy/taxation on the access fee, $t_{i,q} + s_{i,q}$ the net margin (possibly negative) generated on the consumption of household *i* and $c_{i0} + t_{i,q} + s_{i,q}$ the net contribution to service funding of household *i*. Then, as apparent, the comparison of the two Lorenz curves / of the two Gini index is affected (at this stage) by two main effects:

- **the first term** (A_i series) **relates to the impact of subsidies/taxes** on the access fee and consumption which are set up by the IBT for the given consumption levels q_1^{IBT}, q_2^{IBT}, … q_n^{IBT}.

This element is captured by the series of terms A_i which are:

(*i*) positive for households that are net beneficiaries of the cross-subsidy system (with, in this case, $c_{i0} + t_{i,q} + s_{i,q} < 0$ and a household income that increases by $-(c_{i0} + t_{i,q} + s_{i,q}) > 0$ compared to the TBSE);

(*ii*) negative for households that are net contributors to the funding of the cross-subsidy system (with, in this case, $c_{i0} + t_{i,q} + s_{i,q} > 0$ and a household income that decreases by $c_{i0} + t_{i,q} + s_{i,q} > 0$ compared to the TBSE).

This translates into a change in the Lorenz curve of household income net of the water bill and, as a result, into a change in the Gini index.

- **the second term** (B_i series) **relates to the impact of variations in consumption**, linked to the switch from a TBSE to an IBT (potentially financially unbalanced, see below), which per se generate a change in water bills and, therefore, a change in the net incomes $R_i^{net} = R_i - T_i$ of the household, excluding subsidies/taxes on water consumption and access fee.

In particular and as apparent:

- the B_i effect will result in an increase in the net income of household *i* (via lower bills) if the introduction of IBT generates a fall in consumption for household *i* with $q_i^{\text{IBT}} < q_i^{\text{TBSE}}$, i.e. if IBT pricing is water-saving incentive for household *i*, and a fall in the net income of household *i* in the opposite case ($q_i^{\text{IBT}} > q_i^{\text{TBSE}}$)[37].

Taking these two effects into account, the A_i series and the B_i series, leads (*i)* to a comparison of 3 Lorenz curves:

- the Lorenz curve associated with a TBSE with TBSE consumption q_i^{TBSE} (already considered in step 1),
- the Lorenz curve associated with a TBSE with IBT consumption q_i^{IBT} (this Lorenz curve is a new one),
- the Lorenz curve associated with an IBT with IBT consumption q_i^{IBT} (like the first one, it was already entered in step 1),

and (*ii*) to decompose the change in the Gini index following the transition:

$$L_{\text{TBSE}}(q_i^{\text{TBSE}}) \rightarrow L_{\text{IBT}}(q_i^{\text{IBT}}) \quad (9.19)$$

into the sum of two terms referring to the passages:

$$L_{\text{TBSE}}(q_i^{\text{TBSE}}) \rightarrow L_{\text{TBSE}}(q_i^{\text{IBT}}) \rightarrow L_{\text{IBT}}(q_i^{\text{IBT}}) \quad (9.20)$$

The effect on the Gini index linked to the transition:

$$L_{\text{TBSE}}(q_i^{\text{TBSE}}) \rightarrow L_{\text{TBSE}}(q_i^{\text{IBT}}) \quad (9.21)$$

represents an effect on income inequalities linked to the incentive effect of the IBT which the terms $B_i = c \times (q_i^{\text{TBSE}} - q_i^{\text{IBT}})$ are accounting for. The terms $A_i = -(c_{i0} + t_{i,q} + s_{i,q})$ refer to a comparison between the IBT and the TBSE for a given IBT consumption. They relate to the transition:

$$L_{\text{TBSE}}(q_i^{\text{IBT}}) \rightarrow L_{\text{IBT}}(q_i^{\text{IBT}}) \quad (9.22)$$

and refer directly to the impact of subsidies/taxations (Access Fee Included) on income inequalities. The tool then calculates the variation in the corresponding Gini index by considering the Lorenz curve generated by IBT, L_{IBT}, and the Lorenz curve generated by IBT consumption invoiced at a unit price of c and an access fee of $\frac{CF}{n}$, that is $L_{\text{TBSE}}(q_{\text{IBT}})$. A first decomposition of the Gini index is thus obtained with:

$$\Delta i = \Delta_B i + \Delta_A i \quad (9.23)$$

[37] As a reminder, compared to TBSE, the introduction of a social incentive type IBT results in (*i)* an increase in consumption for "small" consumers and (*ii*) a decrease in consumption for "large" consumers with (*iii*) an impact on overall consumption which is then undetermined.

Step 3 One difficulty in comparing the 2 Lorenz curves, $L_{\mathrm{TBSE}}(q_i^{\mathrm{IBT}})$ and $L_{\mathrm{IBT}}(q_i^{\mathrm{IBT}})$, is that the $L_{\mathrm{IBT}}(q_i^{\mathrm{IBT}})$ Lorenz curve is accompanied by an Operating Profit (REX) which is a priori non-zero, and constitutes therefore a direct subsidy by the operator for the benefit of household if it is negative, a direct taxation by the operator against the household if it is negative. In contrast, the $L_{\mathrm{TBSE}}(q_i^{\mathrm{IBT}})$ curve has a REX equal to 0 (by construction). In order to highlight the impact of this REX (and of the potential direct subsidy/taxation on the part of the operator) and identify a net effect of IBT, it is to consider a particular IBT tariff in which:

- the subscription fee sets still to F ;
- all prices π_j are increased by $x > 0$ euros when the REX is negative, and reduced by x euros when the REX is positive, so as to restore the financial equilibrium of the service for the unchanged consumptions $q_i = q_i^{\mathrm{IBT}}$, that is for the consumption levels that are in line with the IBT considered by the user.

In what follows, this reference IBT will be referred to as the Apparently Balanced IBT (IBT-AE)[38]. As this price variation $\Delta\pi_j = x$, for *j* varying from 1 to 4 in the case of an IBT4, does not modify Nordin's D, the equation to be solved writes down:

$$\sum_{i=1}^{n}\left[F - D_i + (\pi_i + x)\,q_i^{\mathrm{IBT}}\right] = CF + c \times \sum_{i=1}^{n} q_i^{\mathrm{IBT}} \tag{9.24}$$

Its solution is:

$$x = \frac{1}{\overline{q}^{\mathrm{IBT}}} \times \left[\left(F - \frac{CF}{n}\right) - \frac{1}{n}\sum_{i=1}^{n}\left[(\pi_i - c)\,q_i^{\mathrm{IBT}} - D_i\right]\right] = -\frac{c_{i0} + \overline{t_q + s_q}}{\overline{q}_i^{\mathrm{IBT}}} \tag{9.25}$$

(see Appendix 6 for details of the calculation) with $\overline{q}_i^{\mathrm{IBT}}$ the average household consumption for the IBT tariff which is assessed/tested by the user, $c_{i0} = F - \frac{CF}{n}$ the subsidy/taxation on the access fee and:

$$\overline{t_q + s_q} = \frac{1}{n}\sum_{i=1}^{n}\left(t_{iq} + s_{i,q}\right) \tag{9.26}$$

the average net contribution, through consumption, of household subscribers to the financing of the system. As apparent, $x = 0$ when the IBT is financially balanced with:

$$\overline{t_q + s_q} = \frac{1}{n}\sum_{i=1}^{n}\left(t_{iq} + s_{i,q}\right) = -c_{i0} = -\left(F - \frac{CF}{n}\right) > 0 \tag{9.27}$$

when the access fee is subsidised ($F < \frac{CF}{n}$).

[38] "Apparently" because this revision of the unit price scale results in a fall in consumption when $x > 0$, an increase in consumption when $x < 0$ (this revision does not change Nordin's D and generates a variation in the marginal price, from π_i to $\pi_i' = \pi_i + x$). This mechanism can be seen as a reimbursement / a contribution to the financing of the service, which would be made ex post and based on the household's water consumption. Other assumptions concerning the financial rebalancing are also possible (for example, the REX could be divided equally with a lump sum of $-\mathrm{REX}/n$ for each household).

With this IBT-AE tariff scheme, the variation A_i in the net income of household *i*:

$$A_i = R_i^{\text{net\_IBT}}(q_i^{\text{IBT}}) - R_i^{\text{net\_TBSE}}(q_i^{\text{IBT}}) = -\left(F - \frac{CF}{n} + (\pi_i - c) \times q_i^{\text{IBT}} - D_i\right) \quad (9.28)$$

is broken down as follows:

$$\begin{aligned} A_i &= \left[R_i^{\text{net\_IBT}}(q_i^{\text{IBT}}) - R_i^{\text{net\_IBT-AE}}(q_i^{\text{IBT}})\right] + \left[R_i^{\text{net\_IBT-AE}}(q_i^{\text{IBT}}) - R_i^{\text{net\_TBSE}}(q_i^{\text{IBT}})\right] \\ &= x \times q_i^{\text{IBT}} - \left[\left(F - \frac{CF}{n}\right) + (\pi_i + x - c) \times q_i^{\text{IBT}} - D_i\right] = A_i' + A_i'' \end{aligned} \quad (9.29)$$

with:

$$A_i' = R_i^{\text{net\_IBT}}(q_i^{\text{IBT}}) - R_i^{\text{net\_IBT-AE}}(q_i^{\text{IBT}}) = x \times q_i^{\text{IBT}} = -\left(c_{i0} + \overline{t_q + s_q}\right) \times \frac{q_i^{\text{IBT}}}{\overline{q}_i^{\text{IBT}}} \quad (9.30)$$

the impact of the REX (under the refinancing method envisaged) on the income, net of the water bill, of household *i* and:

$$A_i'' = R_i^{\text{net\_IBT-AE}}(q_i^{\text{IBT}}) - R_i^{\text{net\_TBSE}}(q_i^{\text{IBT}}) = \left(F - \frac{CF}{n}\right) + (\pi_i + x - c) \times q_i^{\text{IBT}} - D_i \quad (9.31)$$

the impact of the revised IBT (as described above) on the income, net of the water bill, of household *i*. As is apparent, there are therefore 2 sources of variation in household income:

- the first (*cf.* the term A_i') is an ex-post contribution which measures the impact of REX (direct support / direct taxation by the operator) on the Lorenz curve of net income, after payment of the water bill;
- the second (*cf.* the term A_i'') measures the effects of the IBT for an apparently balanced financing of the service, with an ex-post refinancing, based on the consumptions, at a rate of x euros.

In the event of a deficit, the catching up $x > 0$:

(*i*) reduces the amount of the gross subsidy $(\pi_i + x - c)q_i^{\text{IBT}}$ (calculated using a Nordin approach) when $\pi_i < c$, i.e. when the household's consumption is located in a subsidised block,

and:

(*ii*) increases the amount of the gross margin $(\pi_i + x - c)q_i^{\text{IBT}}$ (same remark) when $\pi_i > c$, i.e. when the household's consumption is located in a taxed consumption block.

This effect leads to:

- determine the Lorenz curve associated with an IBT-AE, with IBT consumption q_i^{IBT} and a tariff in which all the unit prices π_j vary by *x* euros (this Lorenz curve is denoted in the following $L_{\text{IBT-AE}}(q_i^{\text{IBT}})$);

- to compare this Lorenz IBT-AE curve with the one that emerges with a TBSE, for the given level of consumptions $q_i = q_i^{\text{IBT}}$, to measure the impact of the structuring of the pricing policy with an apparently balanced financing of the service (i.e. without direct support);
- to calculate the resulting change in the Gini index.

Besides:

- the comparison of the IBT-AE Lorenz curve with the IBT Lorenz curve shows the effect generated by the subsidy/direct taxation by the service operator;
- the resulting change in the Gini index then reflects the impact of this direct subsidy/taxation on income inequality.

Ultimately, the move:

$$L_{\text{TBSE}}(q_i^{\text{TBSE}}) \rightarrow L_{\text{TBSE}}(q_i^{\text{IBT}}) \rightarrow L_{\text{IBT}}(q_i^{\text{IBT}}) \quad (9.32)$$

is therefore broken down into:

$$L_{\text{TBSE}}(q_i^{\text{TBSE}}) \rightarrow L_{\text{TBSE}}(q_i^{\text{IBT}}) \rightarrow L_{\text{IBT-AE}}(q_i^{\text{IBT}}) \rightarrow L_{\text{IBT}}(q_i^{\text{IBT}}) \quad (9.33)$$

with a (related) variation in the Gini index, which in turn can be broken down as:

$$\Delta i = \Delta_B i + \Delta_A i = \Delta_B i + \Delta_{A''} i + \Delta_{A'} i \quad (9.34)$$

To conclude, it should be noted that, with the construction of the Lorenz curve $L_{\text{IBT-AE}}$, there is no change in the subsidy/taxation on the access fee or in the degree of progressivity of the tariff as measured by Nordin's D (price variations $d\pi_j = x$ are simply reflected by a shift up the unit price scale in the case of a deficit, and down in the case of a surplus). In this context, the shift from the IBT-AE Lorenz curve to the IBT Lorenz curve is due to the existence of this direct subsidy/taxation on the part of the service operator when the IBT (considered by the user) is not financially balanced. At the same time, and in all cases, this effect (which must be taken into account for reasoning all other things being equal) is expected to be weak since, given the "water pays for water" principle, the IBT which is evaluated/tested by the user must be financially balanced (at the very least, close to financial equilibrium), with a profit (operating result) equal to (close to) 0. Last, this decomposition of the variation in the Gini index is also implemented with the concentration curve of household income, net of subscriptions, in relation to the standard of living of households and the calculation of the corresponding concentration coefficients.

Illustration numérique Voir Table 33, page 126, for a numerical illustration (with values of the Schutz coefficient also shown). The figure of -0.3 (last figure of the second row) corresponds to the final variation $i_{\text{IBT}} - i_{\text{TBSE}}$ of the Gini index which is then decomposed into the sum of its three components: Incentive Effect; Cross-subsidy System; Direct Support/Taxation. Finally, this decomposition is also displayed with the concentration curve of household income, net of the access fee, in relation to the standard of living of households and the calculation of the corresponding concentration coefficients[39]. On this particular point, see Table 34.

[39] Another possible approach is to think in terms of the population of Individuals and in terms of per capita income (or standard of living). On this point, *see* Dahan & Nisan [2013]. In this version of the tool, this processing is left up to the user (who still has the option of exporting his data for further processing).

	TBSE Income	TBSE-IBT Income	IBT-AE income	-IBT Income
Gini	49.4	49.3	49.1	49.1
Delta	-0.1	-0.2	0.0	-0.3
Schutz	36.3	36.2	36.1	36.1

Table 33 : Change in Gini index (from TBSE to IBT)

	TBSE Income	TBSE-IBT Income	IBT-AE Income	-IBT Income
Quasi-Gini	47.3	47.2	47.1	47.1
Delta	-0.1	-0.1	0.0	-0.2

Table 34 : change in quasi- Gini index (from to IBT)

9.3 Decompositions

The tool completes this equity analysis with:

(i) factor decomposition (decomposition of Rao [1969], decomposition of Pyatt, Chen & Fei [1980]) of the variation of Gini index (of household income net of water bill payment) to assess the impact of gross subsidies, gross contributions to service funding, Incentive effect (*cf*. eq. (9.18)) and direct support (*cf*. eq (9.30)) that are generated by the IBT which is tested / evaluated by the user,

and next:

(ii) with group decomposition following the methodology of Dagum [1997].

In addition to identifying the sources (Inter / Intra / Transvariation) of variation in the Gini index, the decomposition by groups makes it possible, from a Political Economy perspective, to assess the impact of the IBT on the potential polarisation of incomes (with inter-group heterogeneity increasing and intra-group homogeneity increasing).

With additional processing, decomposition by factors (see appendix 7 for a presentation of the methodology) enables to identify the impact of taxation (VAT and environmental charges) and, also, the impact of 'good' and 'bad' subsidies (on basic service and non-basic consumption) and of 'good' and 'bad' 'taxes' (on non-basic consumption and basic service) on household income inequality (and equity, in the sense defined above).

Finally, the tool displays the decomposition by groups and factors (following the methodology of Mussard [2006]) to assess the heterogeneity of the factor impacts by distinguishing two populations: poor households families and non-poor households.

– EVALUATION - ECONOMIC EFFICIENCY

10.1 Foundations

10.1.1 Pareto optimality and aggregate (social) surplus

A) Pareto Criterion The concept of economic efficiency refers to the ability of a system (market economy, centralised management …) to use resources (natural resources, financial resources (capital), human resources (labour factor)) efficiently in view of a minimal criterion known as Pareto optimality. Initially, this criterion simply states that a state of the economy is Pareto optimal if it is not possible to improve the situation of one agent (in this case a household) without at least worsening that of another agent (in this case another household). In the opposite case, i.e. if it is possible to do better for one agent (household) without doing worse for the others, there are possibilities of improving the general situation of the community (i.e. Pareto improvements are possible) and the state of the economy (under consideration) is then referred as sub-optimal (in the Pareto sense). This point is illustrated in Figure 22, page 128, for a simple case in which only 2 household stand in the economy (which makes it possible to draw a graph). On the x-axis is the level of well-being (utility index) of household 1 and, on the y-axis, that of household 2. The set represented here is a Utility Possibility Set (UPS). It is constructed by calculating, for each possible allocation of the economy[40], the utilities of agents 1 and 2, which are indicators of the well-being /of the situation of the agents in each of the states. It is then understood that,

- if the way the system works puts the community at a point such as $\mathrm{A} = (U_1^{\mathrm{A}}, U_2^{\mathrm{A}}) = (0.3, 6)$, the underlying situation/allocation considered is not Pareto optimal, because it is possible to improve the situation of one agent, agent 1 or agent 2, without worsening the situation of the other.

Conversely,

- if the way the economy works puts the community at a point located on the upper decreasing frontier of the UPS, such as the point $\mathrm{B} = (U_1^{\mathrm{B}}, U_2^{\mathrm{B}}) = (0.5, 7.5)$ for instance, the system can be described as efficient in the sense that it makes the most of the possibilities offered in terms of producing collective Welfare.

A feature of the point $\mathrm{B} = (U_1^{\mathrm{B}}, U_2^{\mathrm{B}}) = (0.5, 7.5)$ is that it is not possible to improve the situation of one agent without worsening that of the other. This property makes it possible to qualify point B and the allocation that generates this point B as a Pareto optimum, by construction, and this conclusion holds for all points located on the decreasing upper frontier of the UPS. The set of allocations that generate these Pareto-optimal pairs of utilities are then called social optima because they are socially efficient / located on the upper frontier of the UPS.

[40] An allocation is a list (vector) that specifies an allocation of the resources available to an economy for the production of goods and services and a distribution of these productions, for consumption purposes, between the different agents (in this case household) in the economy. Given the employment-resource constraints, there are a multitude of possible allocations (potentially an infinite number) and each allocation has an impact, through consumption profiles, on the situations (utilities) of household.

Figure 22 : Pareto Efficiency / Pareto Improvement

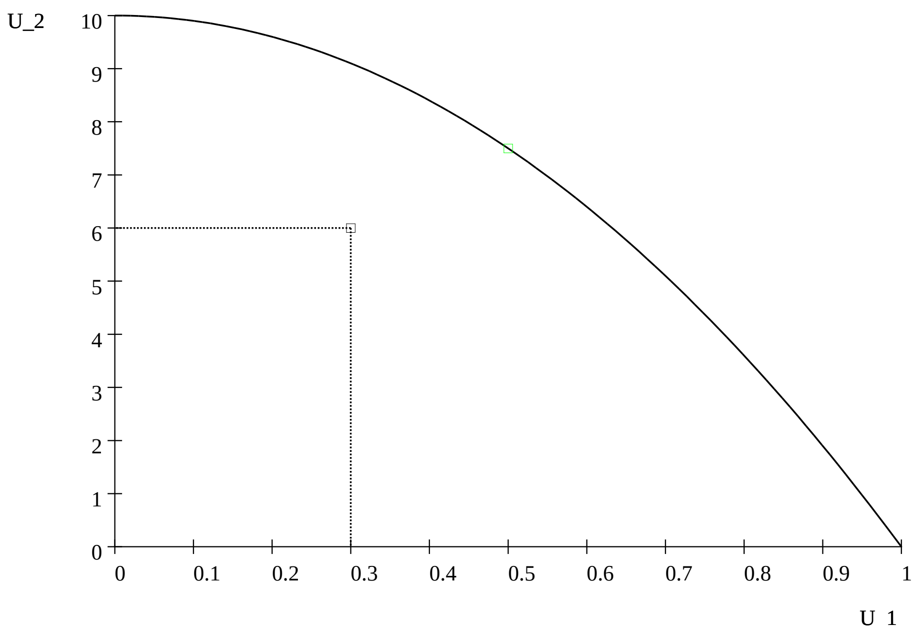


Figure 22-A Pareto Efficiency

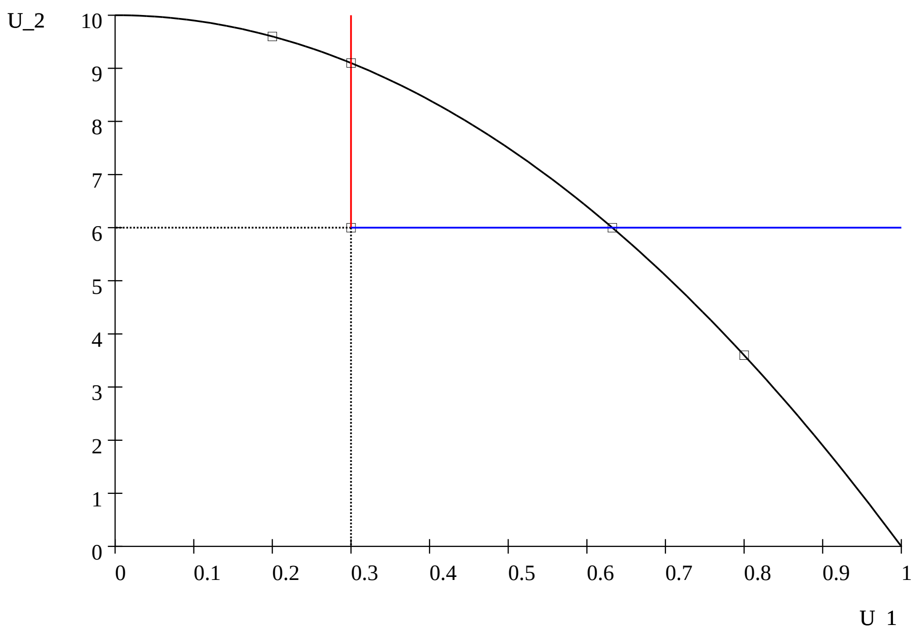


Figure 22-B: Pareto Improvement

As apparent, the Pareto optimality criterion is a minimum efficiency criterion, making it possible to assess the ability of an economic system to use and allocate resources and the production of goods and services satisfactorily with regard to the interests (utilities) of consumers, in this case the households. On the other hand, this same Pareto optimality criterion does not concern itself with redistributive aspects, that affect the sharing of the production of goods and services between the various agents in the economy. In particular: (i) a situation in which one agent has everything and the others nothing can constitute an optimum in the Pareto sense; (ii) a Pareto dominated situation may be not dominated, in the Pareto sense, by a Pareto optimal situation. For instance, points $\mathrm{C}=(0.2, 9.6)$ and $\mathrm{D}=(0.8, 3.6)$ in Figure 22B are Pareto optima, but they do not dominate in the Pareto sense point $\mathrm{B}=(0.5, 7.5)$ which is dominated in the Pareto sense. More generally, the question of the distribution of wealth (that is produced by economic activity) is (of course) an important issue, for which there are tools that can be used to assess whether it is more or less equitable (such as the theory of fair allocations, for example), but which is not dealt with by using the concept of Pareto optimality.

B) Aggregate surplus ("*the quasi-linear framework*") This analytical framework can be extended substantially when one is interested in the organisation of the production of a good for which the demand functions of the agents do not depend (or only slightly) on income (in the sense that the income effects of a variation in price are weak, which is empirically the case for household water demand functions). Accompanied by a system of adequate transfers (described below), it is shown[41] that a Pareto optimal (or socially efficient) allocation is determined by maximising the criterion:

$$\Gamma(q_1, q_2, \ldots, q_n) = u_1(q_1) + u_2(q_2) + \ldots + u_n(q_n) - C(q_1 + q_2 + \ldots + q_n) \qquad (10.1)$$

with :

- $u_1(q_1)$, $u_2(q_2)$, ..., $u_n(q_n)$ are some functions known as "gross surplus functions" which give the monetary value for agent *i* of a consumption (in this case) of water equal to q_i, $i = 1, \ldots, n$;
- $C(q_1 + q_2 + \ldots + q_n)$ is the production cost of the service level $Q = q_1 + q_2 + \ldots + q_n$.

This criterion (which should therefore be maximised for the implementation of a socially efficient / Pareto optimal use of the resource) is called the aggregate surplus (or social surplus) and has a clear interpretation. Basically, it measures the difference between :

- what the community of the n agents (in this case, household) are willing to pay to consume (collectively) the consumption profile $(q_1, q_2, \ldots, q_n)$

and :

- how much it costs to provide a service level $Q = q_1 + q_2 + \ldots + q_n$ to meet this consumption profile $(q_1, q_2, \ldots, q_n)$.

In this way, the aggregate surplus is akin to a kind of social profit.

[41] See for instance Mas Colell et al [1995].

The aggregate surplus approach is therefore the basis for cost-benefit analysys and is prima facie a reference framework for public action. Figure 23, page 131, give the main features of this evaluation framework (in the case of $n = 2$ households, to illustrate).

The first element is that, when demand functions do not depend (or only slightly) on income, the utility functions of agents $U_i(q_i, q_{2i})$ (defined on water consumption and the composite good "Other goods", as in section 5.1) admit (or can be approximated by) so-called quasi-linear forms:

$$U_i(q_i, q_{2i}) = u_i(q_i) + q_{2i} = u(q_i) + R_i - T_i \tag{10.2}$$

with q_i the water consumption of household *i*, R_i its income, T_i the amount of a (lump-sum) tax levied by Public Authorities to finance the EP / EPA service and $q_{2i} = R_i - T_i$ the consumption (calculated by balance) of the other goods ("composite good") of household *i*. The function $u_i(\cdot)$ designates the gross surplus function of household *i* (*cf.* equation (10.1) and the definition of aggregate surplus). The point is then that these indicators of Household Well-Being (utility) are linear in euros and, as a result, any monetary transfer between agents is equivalent to a transfer of Well-Being (utility) of the same amount. Within this framework, the problem of the public decision-maker (planner) can be presented as follows.

(1) Going back to the construction of the UPS, the implementation of allocation A:

$$\mathrm{A} = \left(q_1^{\mathrm{A}}, q_2^{\mathrm{A}}, q_{21}^{\mathrm{A}}, q_{22}^{\mathrm{A}}, x, Q^{\mathrm{A}}\right) = \left(q_1^{\mathrm{A}}, q_2^{\mathrm{A}}, R_1 - T_1^{\mathrm{A}}, R_2 - T_2^{\mathrm{A}}, T_1^{\mathrm{A}} + T_2^{\mathrm{A}}, q_1^{\mathrm{A}} + q_2^{\mathrm{A}}\right) \tag{10.3}$$

still produces a pair of utilities $(U_1^{\mathrm{A}}, U_2^{\mathrm{A}})$ but with:

$$U_1^{\mathrm{A}} = u_1(q_1^{\mathrm{A}}) + q_{21}^{\mathrm{A}} = u_1(q_1^{\mathrm{A}}) + R_1 - T_1^{\mathrm{A}} \tag{10.4}$$

$$U_2^{\mathrm{A}} = u_2(q_2^{\mathrm{A}}) + q_{22}^{\mathrm{A}} = u_2(q_2^{\mathrm{A}}) + R_2 - T_2^{\mathrm{A}} \tag{10.5}$$

where T_1^{A} and T_2^{A} designate the amounts of financial resources of agents 1 and 2 that are allocated to fund the service level $Q_{\mathrm{A}} = q_1^{\mathrm{A}} + q_2^{\mathrm{A}}$.

With a cost function of the form $C(Q) = CF + cQ$ (corresponding to the one implemented in the tool), the sum of these financial contributions set to:

$$T_1^{\mathrm{A}} + T_2^{\mathrm{A}} = C_{\mathrm{A}} = C(Q_{\mathrm{A}}) = CF + c(q_1^{\mathrm{A}} + q_2^{\mathrm{A}}) \tag{10.6}$$

with $C_{\mathrm{A}} = C(Q_{\mathrm{A}})$ the production cost linked to the service level $Q_{\mathrm{A}} = q_1^{\mathrm{A}} + q_2^{\mathrm{A}}$. The point is that any increase in the financial contribution of agent 1, T_1, reduces the financial contribution of agent, T_2, by the same amount and vice versa (for a given level of service $Q_{\mathrm{A}} = q_1^{\mathrm{A}} + q_2^{\mathrm{A}}$ and a fixed consumption plan $(q_1^{\mathrm{A}}, q_2^{\mathrm{A}})$). In so doing,

Figure 23 :

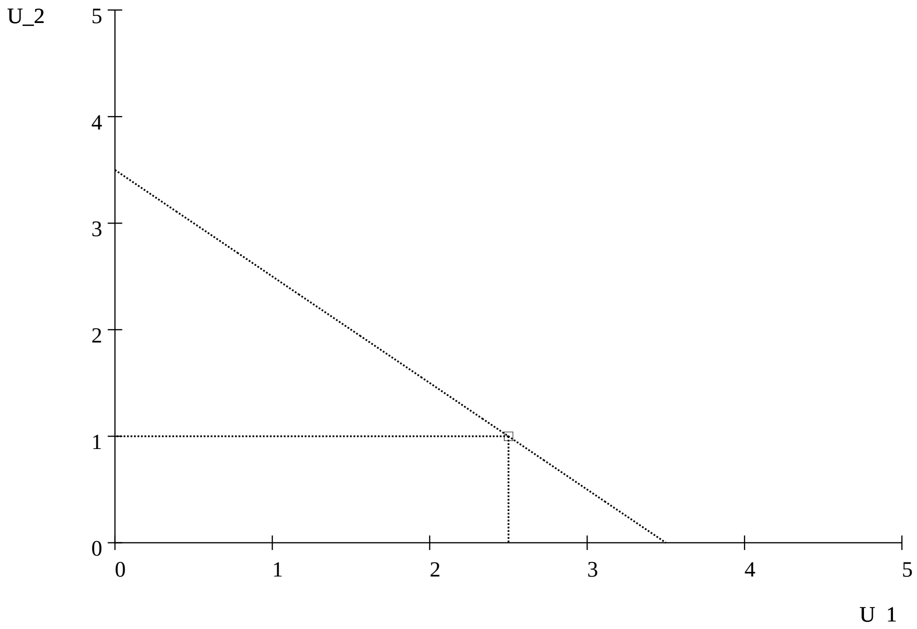


Figure 23-A: Transferable utilities in money for a given A allocation (and $(U_1, U_2) = (U_i^A, U_2^A)$ *)*

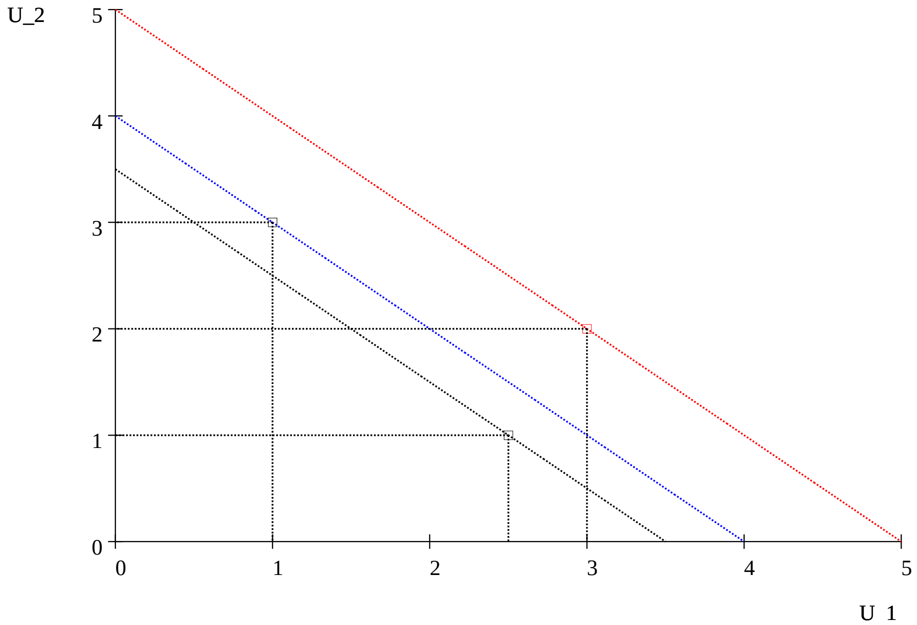


Figure 23-B: Transferable utilities in money with variable allocations A, B, C ... and Pareto Optima

(2) By varying the distribution (T_1, T_2) of the funding $T_1 + T_2 = CF + c(q_1^A + q_2^A)$ relating to the consumption plan (q_1^A, q_2^A) (and the related service level $Q_A = q_1^A + q_2^A$), any point (U_1, U_2) situated on the straight line segment passing through the point $A = (U_1^A, U_2^A)$ and with slope -1 is a feasible pair of utilities (U_1, U_2). Formally:

$$\begin{aligned} U_2 &= u_2(q_2^A) + R_2 - T_2 = u_2(q_2^A) + R_2 - \left[CF + c(q_1^A + q_2^A) - T_1\right] \\ &= u_2(q_2^A) + R_2 - \left[CF + c(q_1^A + q_2^A) - \left(u_1(q_1^A) + R_1 - U_1\right)\right] \\ &= R_1 + R_2 + u_1(q_1^A) + u_2(q_2^A) - \left[CF + c(q_1^A + q_2^A)\right] - U_1 \\ &= c_A^{ste} - U_1 \end{aligned} \qquad (10.7)$$

with :

$$c_A^{ste} = R_1 + R_2 + u_1(q_1^A) + u_2(q_2^A) - \left[CF + c(q_1^A + q_2^A)\right] \qquad (10.8)$$

the y-intercept of the line passing through the point $A = (U_1^A, U_2^A)$, with slope -1, which is effectively a constant for the considered allocation A (*see* equation (10.3)). On these various elements, see Figure 23-A.

Then, the point is that this property applies to all possible feasible allocations:

$$B = \left(q_1^B, q_2^B, q_{21}^B, q_{22}^B, x_B, Q^B\right) = \left(q_1^B, q_2^B, R_1 - T_1^B, R_2 - T_2^B, T_1^B + T_2^B, q_1^B + q_2^B\right) \qquad (10.9)$$

$$C = \left(q_1^C, q_2^C, q_{21}^C, q_{22}^C, x_C, Q^C\right) = \left(q_1^C, q_2^C, R_1 - T_1^C, R_2 - T_2^C, T_1^C + T_2^C, q_1^C + q_2^C\right) \qquad (10.10)$$

...

with associated constants given by:

$$c_B^{ste} = R_1 + R_2 + u_1\left(q_1^B\right) + u_2\left(q_2^B\right) - \left[CF + c\left(q_1^B + q_2^B\right)\right] \qquad (10.11)$$

$$c_C^{ste} = R_1 + R_2 + u_1\left(q_1^C\right) + u_2\left(q_2^C\right) - \left[CF + c\left(q_1^C + q_2^C\right)\right] \qquad (10.12)$$

... (see Figure 23-B). In this context, it is understood that:

(4) Determining the Pareto optimal allocation, taking into account the transfers generated by the distribution (T_1, T_2) of the funding $T_1 + T_2 = C(q_1 + q_2)$, requires to select the pair (q_1, q_2) for which the constant:

$$c^{ste} = R_1 + R_2 + u_1(q_1) + u_2(q_2) - [CF + c(q_1 + q_2)] \qquad (10.13)$$

is maximum, so the pair (q_1, q_2) for which the aggregate surplus:

$$\Gamma(q_1, q_2) = u_1(q_1) + u_2(q_2) - [CF + c(q_1 + q_2)] \qquad (10.14)$$

is maximum.

Last and noting $\left(q_1^*, q_2^*\right)$ this consumption profile in which the aggregate surplus / the constant (10.13) takes its maximum value (we return to the characterisation of this consumption profile later):

(5) There are an infinite number of Pareto optima, all of which are allocations:

$$\mathrm{P} = \left(q_1^*, q_2^*, R_1 - T_1, R_2 - T_2, T_1 + T_2, q_1^* + q_2^*\right) \tag{10.15}$$

for which $T_1 + T_2 = C\left(q_1^* + q_2^*\right)$ with:

(6) the upper boundary of the UPS that consists of all pairs (U_1, U_2) for which:

$$U_2 = R_1 + R_2 + \Gamma\left(q_1^*, q_2^*\right) - U_1 \tag{10.16}$$

(represented by the red line in Figure 23-B).

In this context:

(7) the service level $Q = q_1 + q_2$ is well defined, with $Q^* = q_1^* + q_2^*$

and:

(8) the setting of the contributions (T_1, T_2) to fund the service cost $C\left(Q^*\right) = CF + c\left(q_1^* + q_2^*\right)$ leads to the selection of a specific point on the upper boundary of the UPS.

C) Measuring social inefficiency This approach also makes it possible to measure the potential inefficiency of a system in a simple way. Thus, an organisation that produces, through the allocation A that it puts in place, a pair of utilities $(U_1^{\mathrm{A}}, U_2^{\mathrm{A}})$ for which the aggregate surplus:

$$\Gamma(q_1^{\mathrm{A}}, q_2^{\mathrm{A}}) = u_1(q_1^{\mathrm{A}}) + u_2(q_2^{\mathrm{A}}) - C(q_1^{\mathrm{A}} + q_2^{\mathrm{A}}) = u_1(q_1^{\mathrm{A}}) + u_2(q_2^{\mathrm{A}}) - \left[CF + c(q_1^{\mathrm{A}} + q_2^{\mathrm{A}})\right]$$

is not maximal leads to a loss of social efficiency that can be measured by:

$$\Delta\Gamma = \max_{q_1, q_2} \Gamma(q_1, q_2) - \Gamma(q_1^{\mathrm{A}}, q_2^{\mathrm{A}}) = \Gamma^* - \Gamma(q_1^{\mathrm{A}}, q_2^{\mathrm{A}}) \tag{10.17}$$

with $\Gamma^* = \Gamma\left(q_1^*, q_2^*\right)$ the maximum value of the aggregate surplus. Geometrically, this loss corresponds simply to the vertical distance separating point $(U_1^{\mathrm{A}}, U_2^{\mathrm{A}})$ from the upper boundary of the UPS generated by the level of production Q and the related distribution $(q_1, q_2, Q) = (q_1, q_2, q_1 + q_2)$ that maximises the aggregate surplus $\Gamma = \Gamma\left(q_1, q_2\right)$. Figure 24 illustrates this property.

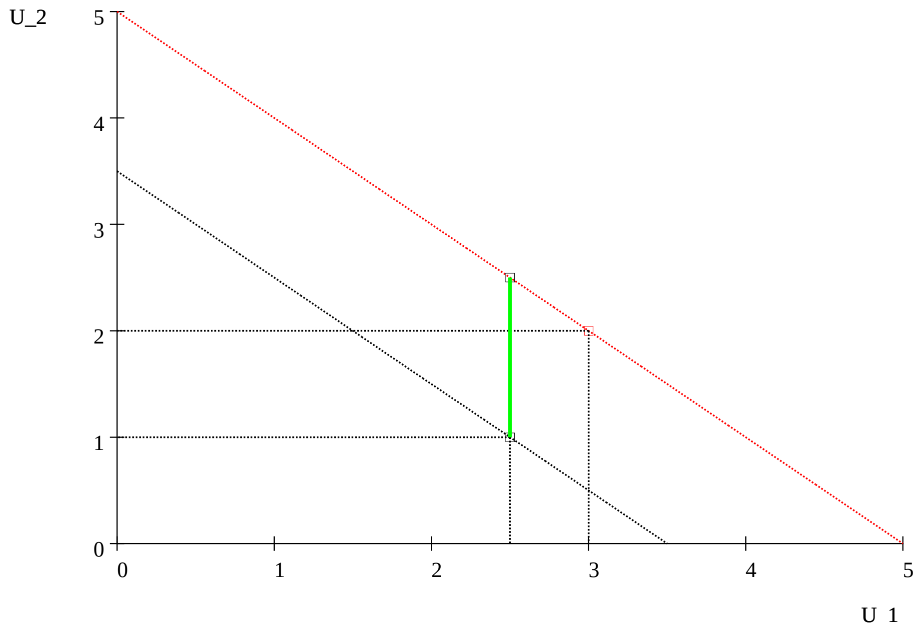


Figure 24 : Measuring Paretian inefficiency - loss of aggregate (social) surplus

One will conclude by emphasising that the aggregate surplus/quasi-linear framework is a veritable paradigm for the conception of public action (the image of the production and distribution of a cake is often used in this respect). Initially, it leads to consider the existence of two departments with:

- an Economic Efficiency department which focusses on deploying the necessary incentive mechanisms to ensure that the system implements an allocation maximizing the aggregate surplus, i.e. focus on maximising the size of the 'cake';
- a Distribution department which, on the basis of this optimal policy, would implement a socially optimal transfer policy $(T_1, T_2, \ldots, T_n)$, i.e. would decide how to divide up this 'maximum-sized cake'.

It should therefore be borne in mind that any public policy approach based on a Cost-Benefit analysis of the Aggregate Surplus type is part of this organisation of public action and, therefore, assumes that the appropriate compensatory mechanisms, through the transfer policy, can be implemented. In particular, it is quite possible that, starting from a socially inefficient situation, the implementation of the Pareto Optimum be spontaneously accompanied by a loss for an agent (under the "non-satiation" hypothesis, it suffices that agent's consumption in the initial situation be greater than the one specified by the Pareto Optimal allocation for this to happen). The point is that the possibility of transfers makes it possible to compensate the agent for this loss, and this aspect of the system, which consists of "paying to reform", is essential for securing the Paretian improvements that are generated by the implementation of the Pareto optimum. However, public policy regulations mean that such compensatory transfers are not always possible and, when this is the case, one may want to stick with a Pareto-dominated allocation, so as not to penalise an agent or category of agents who, in the absence of an adequate compensatory mechanism, would ultimately lose out on the implementation of a Pareto optimum.

Despite its limitations, the maximisation of the aggregate surplus nevertheless constitutes a reference hypothesis for the analysis of public action and the objective function that can be attributed to a public company (in the same way that the maximisation of profit constitutes a reference hypothesis for the objective function of a private company). It is for this reason that this Paretian efficiency item is used to assess the (multi-dimensional) performance of the water and wastewater pricing policy.

10.1.2 First-best allocation

10.1.2.1 Characterisation of the First-Best Allocation

Determining the social optimum, understood as maximising aggregate surplus, poses no particular problem. Given the assumptions about the cost function, the latter is realised as an allocation $(q_1, q_2, \ldots, q_n)$ in which:

$$u_1'(q_1) = u_2'(q_2) = \ldots = u_n'(q_n) = c \qquad (10.18)$$

with:

- $u_i'(q_i)$ the derivative of the gross surplus function of agent *i* with respect to its consumption q_i,
- c the unit variable cost / marginal cost of production (assumed to be constant).

The variable $u_i'(q_i)$ is then given a precise interpretation: it measures **the marginal willingness to pay for water** of agent *i*, that is:

- the amount he is prepared (willing) to pay to increase his water consumption by one unit, starting from the consumption level q_i.

This marginal willingness $u_i'(q_i)$ is then assumed to fall with the level of consumption q_i of agent *i*, i.e. the more agent *i* consumes, the less he is prepared to pay to increase his consumption by one unit. Solving for q_1, q_2, ..., q_n the condition (10.18) gives the so-called first-order allocation:

$$q^* = \left(q_1^*, q_2^*, \ldots, q_n^*\right) \qquad (10.19)$$

with a service level ('first best' or Pareto optimal) which sets to:

$$Q^* = \sum_{i=1}^{n} q_i^* \qquad (10.20)$$

The social optimum is thus characterised by the equality of households' marginal willingness to pay for water (a condition that will not be meet with the application of IBT, what will be a source of inefficiency). See the box on the next page for an illustration in a simple case.

Box - the linear case Figure 25 illustrates the determination of the first-best optimum in a simple case of $n = 2$ households with linear marginal willingness-to-pay of the form:

$$u_i'(q_i) = \alpha_i - \beta q_i$$

where α_1, $\alpha_2 > \alpha_1$ and β are preference parameters (and gross surplus functions of the form $u_i(q_i) = \alpha_i q_i - \frac{\beta}{2} q_i^2$). Pareto-optimal consumption is then given by:

$$u_1'(q_1) = \alpha_1 - \beta q_1 = c \iff q_1^* = \frac{\alpha_1 - c}{\beta}$$

$$u_2'(q_2) = \alpha_2 - \beta q_2 = c \iff q_2^* = \frac{\alpha_2 - c}{\beta}$$

and the socially efficient level of service by:

$$Q^* = q_1^* + q_2^* = \frac{\alpha_1 - c}{\beta} + \frac{\alpha_2 - c}{\beta}$$

The same procedure is then used in the case of *n* agents with gross surplus functions $u_i(q_i)$ that are usually assumed to be increasing ($u_i'(q_i) > 0$) and concave ($u_i''(q_i) < 0$).

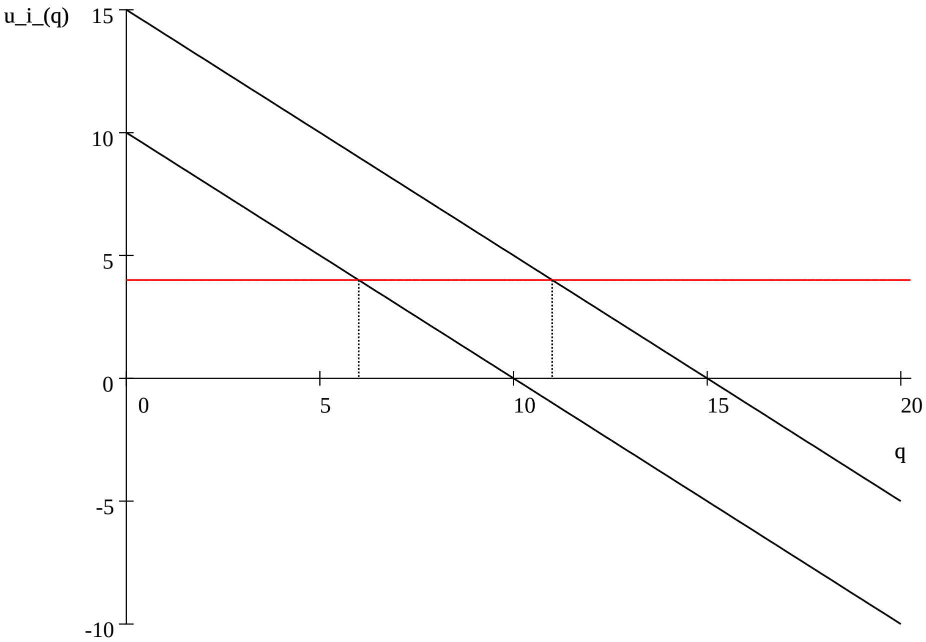


Figure 25 : Determining the first-best optimum

End Box

10.1.2.2 Decentralisation of the First-Best Allocation

The first-order allocation $(q_1^*, q_2^*, \ldots, q_n^*)$ defined by the condition (10.18) can be implemented in a decentralised way by pricing at marginal cost, i.e. by applying :

- linear pricing in which $\pi = c$

or :

- two-part tariff of parameters $(F, \pi) = (\frac{CF}{n}, c)$, that is the “TBSE” scheme.

Indeed, with quasi-linear preferences, determination of the household's optimum consumption (its demand function more generally) is done simply by equating its marginal willingness to pay for water (which is assumed to depend little or not on its income), $u_i'(q_i)$, to the marginal cost of its consumption $T'(q_i)$ which has the particularity to be flat (constant) under these two pricing schemes[42]. It should also be noted that, in the context of an analysis of aggregate surplus (that is, with the “quasi-linear preferences” framework), the choice of the pricing system (marginal cost pricing vs TBSE) is of little importance in terms of efficiency. In particular, marginal cost pricing $\pi = c$ generates an operating deficit $\Delta\Pi = -CF$ which is then ipso facto financed by adjusting, in one way or another, the vector of the financial contributions $(T_1, T_2, \ldots, T_n)$ that are levied to fund the level of service Q^* through, in particular, the organisation of the tax system (on income). In practice, the mobilisation of public funds (to fund an operating deficit) is tainted by an organisational cost which, when explicitly considered, leads to the choice of the TBSE scheme (and the principle "water pays for water").

10.1.2.3 Taking environmental costs into account

In presence of pollution, marginal cost pricing must be understood in the full sense of the term, i.e. including the environmental cost c_e (and the principle "water must be returned clean to nature"). Figure 26, page 138, illustrates this point by assuming, for sake of simplicity, that the marginal willingness to pay for water of the agent is linear (*see* the box "The linear case", page 136). At the origin, the first-best optimum consists of the point:

$$\mathrm{P} = \left(q_i^*, c_{\mathrm{EP}} + c_e\right) = \left(\frac{\alpha_i - (c_{\mathrm{EP}} + c_e)}{\beta}, c_{\mathrm{EP}} + c_e\right) \qquad (10.21)$$

(P for Pareto) in which an agent's marginal willingness to pay for water, $u_i'(q_i) = \alpha_i - \beta q_i$, is equal to the marginal social cost $c_{\mathrm{EP}} + c_e$. The contribution to the aggregate surplus Γ^* of user *i*, γ_i^*, is then calculated using the integral:

$$\gamma_i^* = u_i(q_i^*) - (c_{\mathrm{EP}} + c_e)q_i^* = \int_0^{q_i^*} u_i'(t)dt - (c_{\mathrm{EP}} + c_e)q_i^* = \int_0^{q_i^*} \left[u_i'(t) - (c_{\mathrm{EP}} + c_e)\right]dt \qquad (10.22)$$

[42] This principle, according to which the agent determines his consumption by equalising his marginal willingness to pay for water with the unit price of consumption, does not necessarily hold with an IBT (and the case of corner solutions $q_i^* = k_j$; see below).

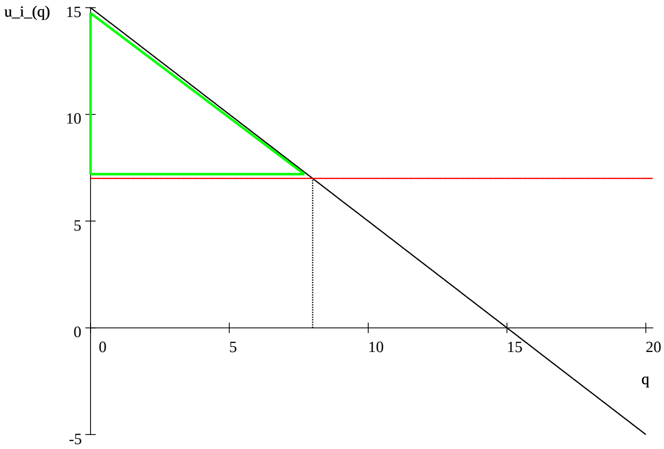


Figure 26 : Contribution to household i's aggregate surplus - first-best optimum[43]

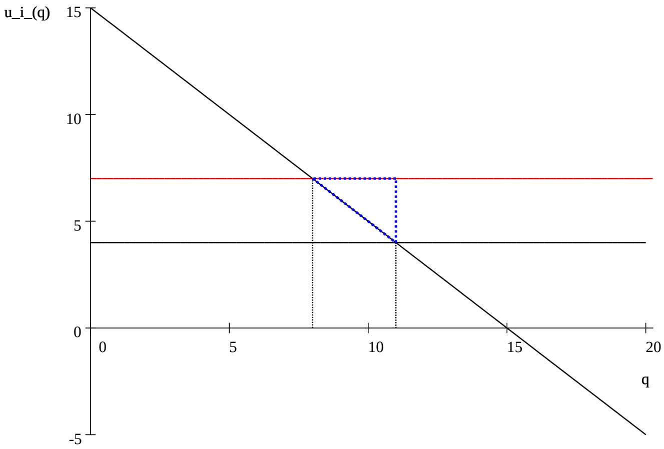


Figure 27 : Aggregate surplus loss on EPA consumption of household i[44]

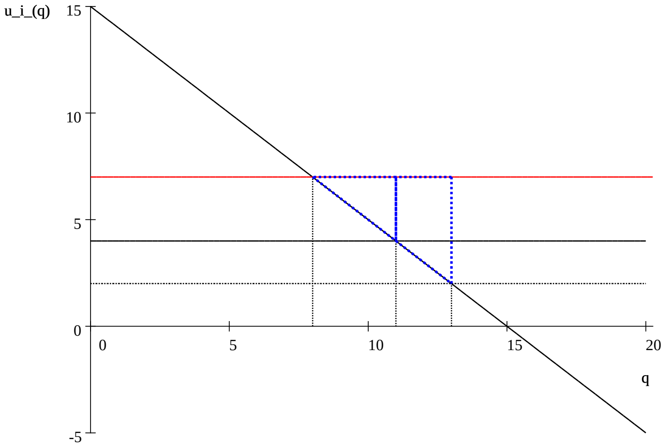


Figure 28 : Loss of aggregate surplus on EP *consumption for a user i*

[43] With full cost recovery.
[44] With incomplete cost recovery.

This area is represented by the area of triangle ABP, shown in green in Figure 26. Then, when cost recovery is incomplete, a subscriber to the drinking water and wastewater service is billed in the TBSE-EPA at a unit price of:

$$\pi_{\text{TBSE-EPA}} = c_{\text{EP}} + r_{\text{EP}} + c_{\text{A}} + r_{\text{A}} < c_{\text{EP}} + c_e \tag{10.23}$$

The result is a higher level of consumption compared with the full-cost TBSE:

$$q_i^{\text{TBSE-EPA}} = \frac{\alpha_i - \left(c_{\text{EP}} + r_{\text{EP}} + c_{\text{A}} + r_{\text{A}}\right)}{\beta} > q_i^* = \frac{\alpha_i - \left(c_{\text{EP}} + c_e\right)}{\beta} \tag{10.24}$$

(*see* the transition from point P to point E in Figure 27) and a loss of aggregate surplus which is calculated at the individual level as:

$$\begin{aligned} \Delta\gamma_i^{\text{TBSE-EPA}} &= \gamma_i^{\text{TBSE-EPA}} - \gamma_i^* \\ &= \int_0^{q_i^{\text{TBSE-EPA}}} \left[u'(t) - (c_{\text{EP}} + c_e)\right]dt - \int_0^{q_i^*} \left[u'(t) - (c_{\text{EP}} + c_e)\right]dt \\ &= \int_{q_i^*}^{q_i^{\text{TBSE-EPA}}} \left[u'(t) - (c_{\text{EP}} + c_e)\right]dt \end{aligned} \tag{10.25}$$

This (oriented) area is represented by the area of the PEC triangle (drawn in blue) in Figure 27 and reflects an inefficiency linked to the fact that the consumption of user *i* is too high compared with the first-order optimum.

Considering the modelling of the environmental cost, the loss for a user connected to the drinking water service only is calculated in a similar way, but with the consumption of household *i* being (all other things being equal) greater:

$$q_i^{\text{TBSE-EP}} = \frac{\alpha_i - \left(c_{\text{EP}} + r_{\text{EP}}\right)}{\beta} > q_i^{\text{TBSE-EPA}} = \frac{\alpha_i - \left(c_{\text{EP}} + r_{\text{EP}} + c_{\text{A}} + r_{\text{A}}\right)}{\beta} \geq q_i^* = \frac{\alpha_i - \left(c_{\text{EP}} + c_e\right)}{\beta} \tag{10.26}$$

The result is a loss of aggregate surplus, calculated at the individual level, which is greater:

$$\begin{aligned} \Delta\gamma_i^{\text{TBSE-EP}} &= \gamma_i^{\text{TBSE-EP}} - \gamma_i^* \\ &= \int_0^{q_i^{\text{TBSE-EP}}} \left[u'(t) - (c_{\text{EP}} + c_e)\right]dt - \int_0^{q_i^*} \left[u'(t) - (c_{\text{EP}} + c_e)\right]dt \\ &= \int_{q_i^*}^{q_i^{\text{TBSE-EP}}} \left[u'(t) - (c_{\text{EP}} + c_e)\right]dt \end{aligned} \tag{10.27}$$

some of which is due to the fact that the user is not connected to the sewerage system (with discharges that are considered to be untreated or very poorly treated, as discussed in paragraph 3.3.3.2) On these two properties, see Figure 28 on previous page.

The tool then calculates these surplus losses, linked to the incomplete nature of cost recovery, for (*i)* each of the households in the Subscriber File (Population module) and (*ii*) at the global / community level with the aggregate surplus loss:

$$\Delta\Gamma_{\text{TBSE}} = \sum_{i=1}^{n_1} \Delta\gamma_i^{\text{TBSE-EPA}} + \sum_{i=1}^{n_2} \Delta\gamma_i^{\text{TBSE-EP}} \tag{10.28}$$

with n_1 is the number of households connected to the collective sewerage network and $n_2 = n - n_1$ is the number of households that are not. However, this calculation is based on an approximation of household water demand functions, which is described in more detail in the following paragraph. At this stage, it should be emphasised (1) that the result obtained is a monetary sum and (2) that this statistic basically measures the vertical distance between the tuple of monetary utilities $(U_1, U_2, \dots, U_n)$ generated by the allocation thus implemented and the upper frontier of the UPS associated with the implementation of the first-best optimum $\left(q_1^*, q_2^*, \dots, q_n^*\right)$, as illustrated in Figure 24 : Measuring Paretian inefficiency - loss of aggregate (social) surplus, page 134.

10.1.3 Implementation

The quasi-linear framework / the application of this criterion that constitutes aggregate surplus $\Gamma(q_1, q_2, \dots, q_n)$ requires stricto sensu that the demand functions of the agents do not depend on income. The use of this analytical framework (to assess the welfare effectiveness of pricing policy) must therefore be seen here as an approximation, which will be all the better if the income effects of a price variation are low, what is empirically the case for household water demand functions.

Following (notably) Porcher [2014] and Nauges & Whittington [2017], the procedure that is implemented for the calculation of surpluses consists in approximating the water demand function of a household *i* by considering that the amount of the subscription F and the values of Nordin's D are small compared to the value of its quarterly income[45] . When the tariff is properly perceived (polar case $\kappa = 1$), the (conditional) water demand functions given by the relation (5.8) on page 59 can then be approximated by[46] :

$$\ln q_i^d \approx \ln q_{i0} + 0.25 \ln R_i - 0.31 \ln \pi_j \Leftrightarrow q_i^d \approx \frac{A_i}{\pi_j^{0.31}} \qquad (10.29)$$

with $A_i = q_{i0} \times R_i^{0.25}$ a (numerical) constant specific to each household in the subscriber file and π_j the price of block *j* (this approximation also applies to two-part tariffs). In this framework, household *i*'s marginal willingness to pay for water is written as:

$$u_i'(q_i) = A_i^{\frac{1}{0.31}} \times q_i^{-\frac{1}{0.31}} = q_{i0}^{\frac{1}{0.31}} \times R_i^{\frac{0.25}{0.31}} \times q_i^{-\frac{1}{0.31}} \qquad (10.30)$$

(iso-elastic form) with a marginal willingness to pay that is then all the greater (1) the larger the captive portion of consumption q_{0i} and (2) the higher the household income R_i, i.e. compared with low incomes, high incomes are prepared to pay more to increase their consumption by one unit, all other things being equal. The social first-best optimum consumption of household *i* is then given by:

---

[45] The errors made in measuring the "virtual" income of block *j* (i.e. corrected for the amount of the subscription fee and the value of Nordin's D for block *j*) are actually quite small. For instance, for the EP tariff of the city of Saint Paul and a household in which only one adult works for a full-time job at the minimum wage, they are set as a percentage of the true value at 0.2%, -0.8%, -2.2% and -2.7% depending on the consumption block (1, 2, 3 and 4). Now, if both adults work for a full time at the minimum wage, these errors set to 0.09%, -0.4%, -1.1% and -1.4% of the true value.

[46] This function is adjusted for the values of price elasticity (here -0.31) and income elasticity (here 0.25) that are entered by the user.

$$u_i'(q_i) = c_{\text{EP}} + c_e \Leftrightarrow q_i^* = \frac{A_i}{(c_{\text{EP}} + c_{\text{e}})^{0.31}} = q_{0i} \times R_i^{0.25} \times (c_{\text{EP}} + c_{\text{e}})^{-0.31} \quad (10.31)$$

and, neglecting at this stage VAT and environmental duties (that will be introduced later), the consumption implemented by the "EPA" TBSE is given by:

$$q_i^{\text{TBSE-EPA}} = \frac{A_i}{(c_{\text{EP}} + c_{\text{A}} + r_{\text{EP}} + r_{\text{A}})^{0.31}} = q_{0i} \times R_i^{0.25} \times (c_{\text{EP}} + c_{\text{A}} + r_{\text{EP}} + r_{\text{A}})^{-0.31} \quad (10.32)$$

(the result is obtained from solving the equation $u_i'(q_i) = c_{\text{EP}} + r_{\text{EP}} + c_{\text{A}} + r_{\text{A}}$) and the consumption implemented by "EP" TBSE is given by:

$$q_i^{\text{TBSE-EP}} = \frac{A_i}{(c_{\text{EP}} + r_{\text{EP}})^{0.31}} = q_{0i} \times R_i^{0.25} \times (c_{\text{EP}} + r_{\text{EP}})^{-0.31} \quad (10.33)$$

(this result is obtained from solving the equation $u_i'(q_i) = c_{\text{EP}} + r_{\text{EP}}$). Applying the formulae (10.25), (10.27) and (10.28) given on page 139, enables to compute the (estimated) aggregate surplus losses at the household level. The latter are then supplemented by the calculation of per capita losses: $\Delta\gamma_i / N_i$ and $\Delta\Gamma / N$. On the basis of the series thus generated (and included in the Invoices module), the tool provides basic statistics (including mean, median, standard deviation, MAPE, Gini index) enabling to assess the distribution of these (first) inefficiencies in terms of consumption (deviations from first-best optimal social values) and (contributions to) aggregate surplus.

Note - Technical point It should be noted that the calculation of the contribution to the aggregate surplus of household *i* whose consumption is equal to q_i ;

$$\gamma_i(q_i) = \int_0^{q_i} (u_i'(t) - c)dt = \int_0^{q_i} u_i'(t)dt - cq_i = u_i(q_i) - u_i(0) - cq_i \quad (10.34)$$

with $u_i(q_i)$ the gross surplus / the valuation by agent *i* of a level of consumption set to q_i is not (here) well defined since the integral:

$$\int_0^{q_i} u_i'(t)dt = \int_0^{q_i} A_i^{\frac{1}{0.31}} t^{-\frac{1}{0.31}} dt = \frac{31 \times A_i^{\frac{100}{31}}}{69} \times \left(\lim_{q \to 0^+} \frac{1}{q^{\frac{69}{31}}} - q_i^{-\frac{69}{31}} \right) \quad (10.35)$$

is divergent / takes a value equal to $+\infty$ (this point also applies to the social first-best consumption $q_i = q_i^*$). Nevertheless, this property is not without meaning since it implies that the agent is spontaneously prepared to pay an infinite sum, therefore his entire fortune, to consume water in quantity q_i compared to a situation where he would have none. At the same time, it does not preclude the calculation of variations in surplus, in this case variations of the individual contribution to aggregate surplus, which are finite. This property also applies to the calculation of variations in consumer surplus (introduced below).

10.2 IBT assessment

Knowing the first-best allocation, the tool calculates the variations in consumptions (compared to first-best optimum level) and in aggregate surplus that are linked to the implementation of the IBT (which is evaluated/tested by the user) in relation to an "effective" TBSE, in terms of the environmental cost recovery information provided by the user.

These calculations (which are initially carried out at household level) can lead to counter-intuitive results, but these are clearly justified. In order to gain a clear understanding of the mechanisms that are at work, one will illustrate the nature of these properties by considering a simple case in which the user would use an IBT2 for which $\pi_1 < c < \pi_2$ (with subsidy in block 1 and taxation in block 2) and initially disregarding the environmental cost. For the purposes of the analysis, one will also distinguish between two cases depending on whether the tariff is well perceived ($\kappa = 1$) or poorly perceived ($\kappa < 1$).

This assessment of economic efficiency of an IBT concludes with the calculation of (variations in) consumer surplus (which differ from variations in the contribution to the aggregate surplus $\Delta\gamma_i$). These latter indicators measure the gains and losses in utility ("well-being"), achieved / borne by household, measured in euros and that are generated by the change in pricing policy (compared to the one that should be implemented to achieve the first-best social optimum; see above). This information makes it possible to identify winners and losers, within the population of households, when moving from a TBSE water pricing system to an IBT water pricing system.

10.2.1 Analysis - the case $\kappa = 1$

This section sets out the calculation of aggregate surplus losses in the reference case where the tariff (as a reminder, an IBT2 with $\pi_1 < c < \pi_2$) is well perceived. Technically, three cases have to be distinguished depending on whether the household consumption takes place within the block 1 ($q_i^{\text{IBT-PP}} < k_1$), within the block 2 ($q_i^{\text{IBT-PP}} > k_1$) or at the block 1 threshold ($q_i^{\text{IBT-PP}} = k_1$).

The case $q_i^{\text{IBT-PP}} < k_1$ Figure 29 shows the case of a small consumer, defined as a household whose consumption at the social first-best optimum is less than k_1 (the tariff threshold for the IBT2 (F, π_1, π_2, k_1)), with therefore:

$$\ln q_i^* \approx \ln A_i - 0.31 \ln c = \ln q_{i0} + 0.25 \ln R_i - 0.31 \ln c \leq \ln k_1$$

$$\Leftrightarrow \ln R_i \leq 4 \times \left(\ln k_1 + 0.31 \ln c - \ln q_{i0} \right)$$

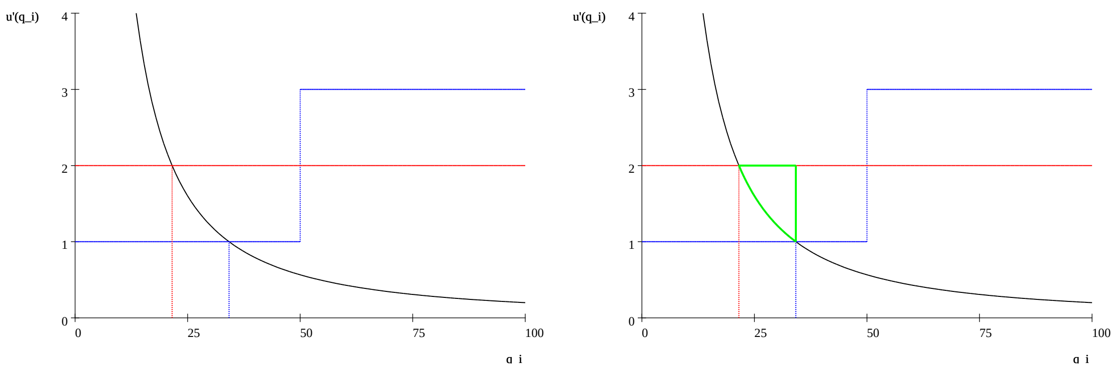
Compared with the TBSE, the implementation of the IBT puts in place a subsidy for the consumption of these small consumers at the rate $\sigma_1 = c - \pi_1$, what leads them (quite naturally) to increase their consumption (*cf.* the shift from point P to point E in Figure 29-A). From a quantitative point of view, the impact is given by:

$$\ln q_i^{\text{IBT-PP}} - \ln q_i^{\text{TBSE}} \approx \left(\ln A_i - 0.31 \ln \pi_1 \right) - \left(\ln A_i - 0.31 \ln c \right) = 0.31 \times \left(\ln c - \ln \pi_1 \right) \quad (10.36)$$

(this last expression gives an approximation of the variation, in %, of the consumption of this agent). This increase in consumption results in a loss of aggregate surplus (in fact, of household *i*'s contribution to the aggregate surplus), which is calculated as:

$$\Delta\gamma_{\text{IBT-PP}} = \int_{q_i^*}^{q_i^{\text{IBT-PP}}} \left[u'(t) - c\right] dt = \int_{q_i^*}^{q_i^{\text{IBT-PP}}} \left(A_i^{\frac{1}{0.31}} \times t^{-\frac{1}{0.31}} - c\right) dt \qquad (10.37)$$

what is represented geometrically by the area of the curvilinear triangle (drawn in green) in Figure 29-B. It should be kept in mind that these efficiency losses took place in the segment of the small consumers, with household consumptions that are too high, compared with their first-best values.



A: determining the optimum for a "small" consumer

B: Loss of aggregate surplus

Figure 29 : The case of a small consumer

The case $q_i^{\text{IBT-PP}} > k_1$ Figure 30 (on next page) shows the impact of IBT2 on the situation of a "large" consumer, defined as an agent whose IBT2 consumption locates in block 2, i.e. for whom:

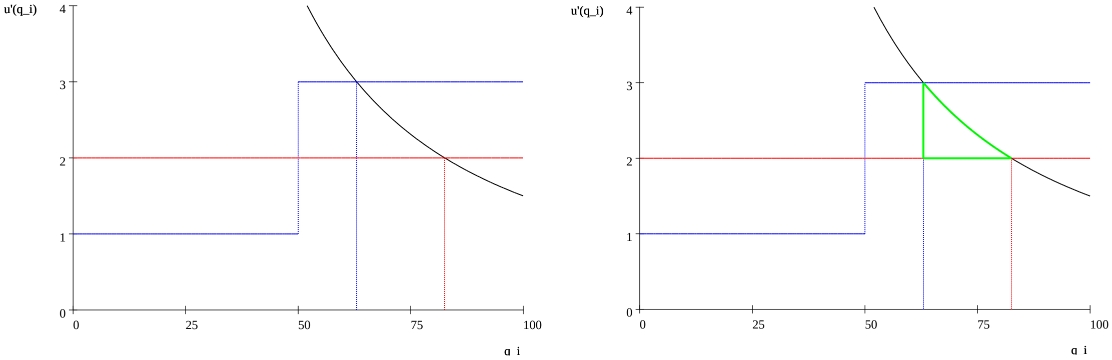
$$\ln q_i^{\text{IBT-PP}} \approx \ln A_i - 0.31 \ln \pi_2 = \ln q_{i0} + 0.25 \ln R_i - 0.31 \ln \pi_2 \geq \ln k_1$$

$$\Leftrightarrow \ln R_i \geq 4 \times \left(\ln k_1 + 0.31 \ln \pi_2 - \ln q_{i0}\right) \equiv \ln \underline{R}_{i2}$$

The point is that, given the need to finance the cross-subsidy system, a "tax" $\tau_2 = \pi_2 - c$ is levied on the consumption of block 2 of this agent who then reduces his consumption, compared to the two-part tariff in which $\pi = c$ (*cf.* the passage from point P_2 to point E_2 on Figure 30-A) with a variation in % in consumption of the order of:

$$\ln q_i^{\text{IBT-PP}} - \ln q_i^{\text{TBSE}} \approx -0.31 \times \left(\ln \pi_2 - \ln c\right) \qquad (10.38)$$

The result is also a fall in aggregate surplus, materialised by the area of the curvilinear triangle (plotted in green) in Figure 30-B, which is quantified by applying the same formula as that given for the case $q_i^{\text{IBT-PP}} < k_1$ (*cf.* the equation (10.37) on page 143). Unlike the previous case, this loss of is the result of consumption which is now too low compared with its first-best social value.



A: determining the optimum for a "large" consumer B: Loss of aggregate surplus

Figure 30 : Loss of aggregate surplus - the case of a "large" consumer

The case $q_i^{\text{IBT-PP}} = k_1$ Figure 31 and Figure 32 (on the next page) show the case of the corner solution $q^d = k_1$ that applies to agents whose level of income verifies:

$$4\times\left(\ln k_1 + 0.31\ln c - \ln q_{i0}\right) \le \ln R_i \le \ln \underline{R}_{i2} = 4\times\left(\ln k_1 + 0.31\ln \pi_2 - \ln q_{i0}\right)$$

In this configuration, the variation in consumption compared with the first-best optimum sets to:

$$\ln q_i^{\text{IBT-PP}} - \ln q_i^{\text{TBSE}} \approx \ln k_1 - \left(\ln A_i - 0.31\times\ln c\right) = \ln k_1 + 0.31\ln c - \ln q_{i0} - 0.25\ln R_i \qquad (10.39)$$

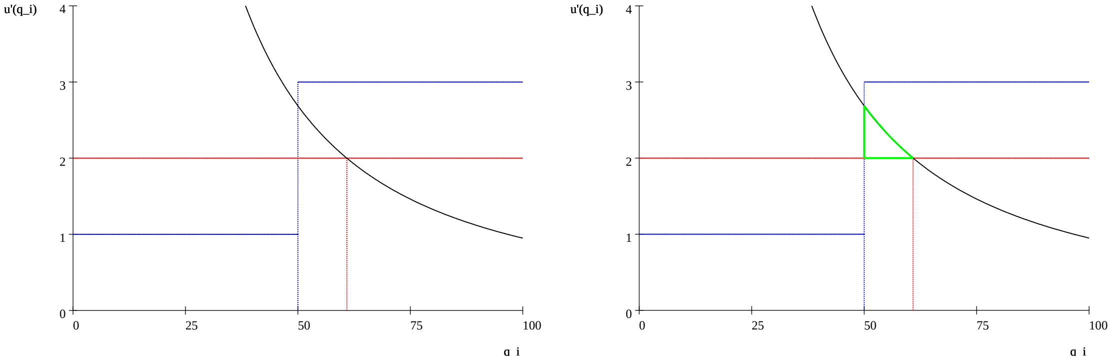
with an impact that can be positive (see Figure 31-A) or negative (see Figure 32-A). In all cases, there is a loss of aggregate surplus (measured by the area of the curvilinear triangles drawn in green in Figure 31-B and Figure 32-B), which is calculated by applying (once again) the formula (10.37).

To sum up, except for the group of households whose characteristics $\left(q_{0i}, R_i\right)$ are such that:

$$\ln q_i^* = \ln A_i - 0.31\ln c = \ln q_{i0} + 0.25\ln R_i - 0.31\ln c = \ln k_1$$

the implementation of the IBT generates losses for the aggregated surplus which are partly linked to consumption that is too high, and partly to consumption that is too low compared with the first-best social values. The tool then sums up all these losses calculated at the household level to calculate the aggregate surplus loss at the aggregate level. For these purposes, it is nevertheless distinguished between 2 categories of users, depending on whether or not the household is connected to the sewerage network, and is also displayed an intermediate step linked to the implementation of an effective TBSE in which households actually connected to the sewerage network face a TBSE-EPA tariff and those who are not face a TBSE-EP tariff.

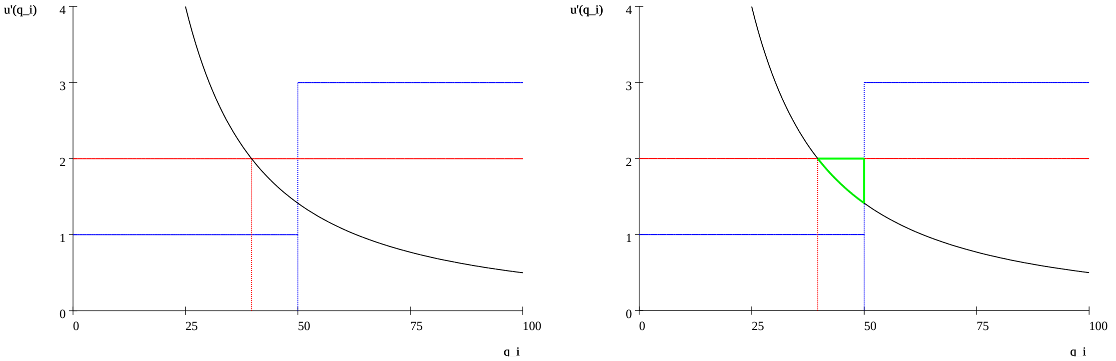
See Crampes & Lozachneur [2014], Mayol A. & Porcher S. [2019] and Paul [2023] for complementary analyses (with, in particular, the endogenization of the price of block 2 and the characterisation of the conditions of financial sustainability of the tariff mechanism) relating to this issue of economic (social) inefficiency, understood in the sense of the aggregate surplus, and therefore of the Pareto Optimality of the IBT scheme (when compensatory monetary transfers are possible at the same time).



A: determination of the optimum

B : Loss of aggregate surplus

Figure 31 : Loss of aggregate surplus – corner solution I



A: determination of the optimum

B : Loss of aggregate surplus

Figure 32 : Loss of aggregate surplus – corner solution II

10.2.2 Analysis - the case $\kappa < 1$

Once these initial elements have been identified, the model adds (to the calculation of household contributions to the aggregate surplus) the over-consumption linked to the poor perception of the tariff, which is measured in the Incentive Effect field of the Evaluation module (it should be noted that, for sake of consistency, the tool does not recalculate this over-consumption on the basis of the approximate demand function (10.29) given on page 140). An important point is that, while this over-consumption is privately costly for agents (see below), its effect on aggregate surplus is complex and can help to reduce the loss of economic efficiency, or even improve Economic Welfare (Paul [2023]) compared to a non full cost recovery TBSE system. This property is consistent with several results highlighted in Tax Economics (see notably Rees-Jones & Taubinsky [2020]).

The Figure 33, page 147, show the point in the case of the IBT2 considered in the previous paragraph with an agent who would reason in terms of average price (the perception parameter κ is then equal to 0). Once remember that the model estimated by BCP does not generate overconsumption for agents located in block 1 (because the average price is calculated excluding subscriptions; see above), it is shown that the poor perception of the tariff for households located in block 2 reduces the negative impact of IBT2 on their contributions to the aggregate surplus (the losses correspond to the areas of the green curvilinear triangles), even when the addition of overconsumption would result in the consumption of household *i* ultimately being higher than its first-best social value (as represented in Figure 33-A, which shows the long-term equilibrium of the dynamic model mentioned in paragraph 5.3 (*see* point (1)). In other words, a poor perception of the tariff is not necessarily adverse from the point of view of aggregate surplus.

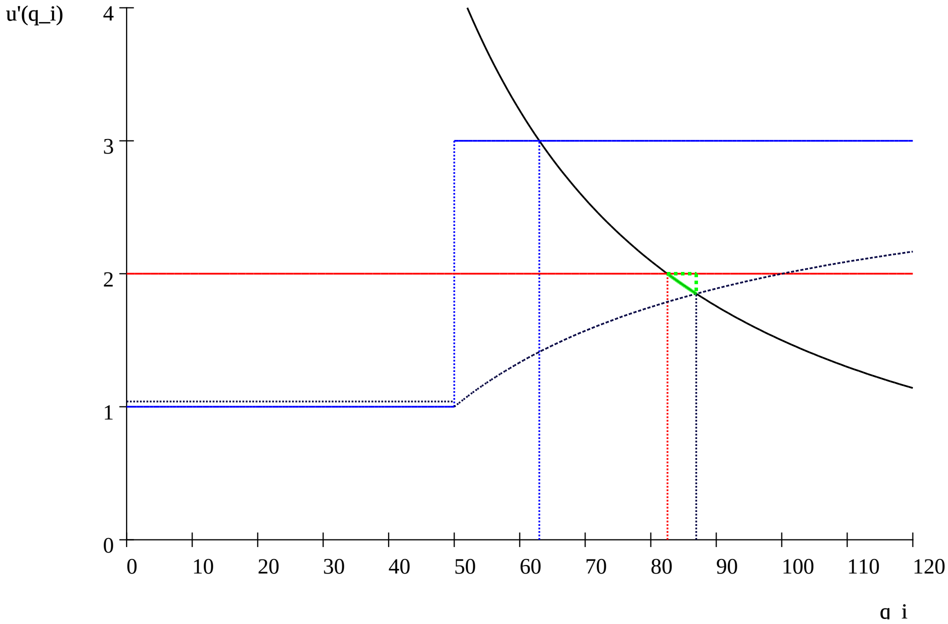
The tool then provides information on the nature of this impact, distinguishing between 2 cases, depending on whether the tariff is properly perceived ($\kappa = 1$) or not (for the value of $\kappa = \kappa_0$ entered by the user), with a breakdown of the impact by consumption block. The addition of overconsumptions (i) has a negative impact on the aggregate surplus in the subsidised blocks from block 2 onwards (or also by leading some households that are initially located in block 1, when $\kappa = 1$, to move up an higher block with $\kappa < 1$) and (ii) tends to reduce the negative impact of the IBT in the consumption blocks that are taxed/subject to contributions to service funding.

Taking into account the environmental cost These elements of analysis presented in the case $\kappa = 1$ continue to apply when cost recovery is complete, with 100% of the subscribers connected to the (collective) sewerage network and a marginal cost of consumption given by $c = c_{\text{EP}} + c_e$. When these conditions are not met, the point is that the implementation of an Increasing Block Tariff may result in a smaller loss of aggregate surplus, compared to a TBSE of parameters $c_{\text{EP}} + r_{\text{EP}} + c_{\text{A}} + r_{\text{A}} < c_{\text{EP}} + c_e$ / when cost recovery is incomplete. This occurs when aggregate IBT consumption decreases (but not too much) compared with aggregate TBSE consumption, with the falls in consumption by large consumers that outweigh the increases in consumption by small consumers (see above). This effect is potentially interesting for the short-term reduction of pollution that cannot be cleaned up / for which the clean-up technology would be too costly in relation to the expected social benefits.

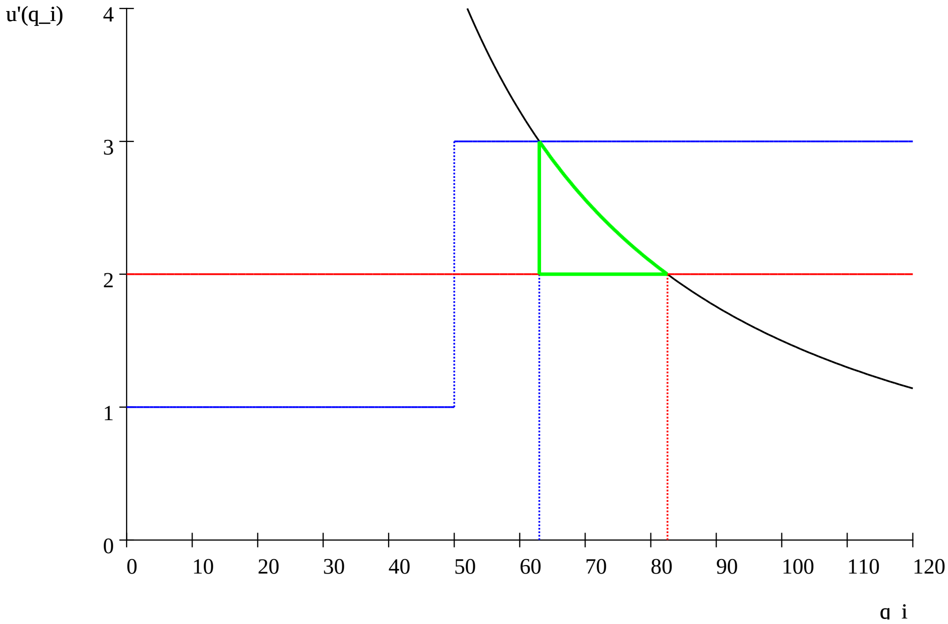
10.2.3 Consumer surplus

The tool completes the calculation of impacts on the aggregate surplus with the calculation of variations in consumer surplus, which are monetary measures of gains and losses in household well-being (they therefore differ from the values taken by household contributions to the social surplus). As certain properties may also appear counter-intuitive, some elements of analysis are provided by distinguishing between 2 main cases depending on whether the perception of the tariff is perfect (case $\kappa = 1$) or imperfect (case $\kappa < 1$). In addition, and as with the impacts on aggregate surplus, the tool provides this information by considering "effective" TBSEs, based on the values entered by the user, as the counterfactual.

Figure 33 : Aggregate surplus losses with $\kappa = 0$ *versus* $\kappa = 1$



A: loss of surplus (aggregated) with $\kappa = 0$



B: Loss of surplus (aggregate) with $\kappa = 1$

10.2.3.1 Tariff is properly perceived ($\kappa = 1$)

Initially, calculating the consumer surplus (or net surplus) requires calculating the quantity:

$$v(q_i) = \int_0^{q_i} u'(t)dt - T(q_i)$$

where (as a reminder) (i) $T(q_i)$ is the amount of the bill that agent *i* has to pay when he consumes the volume of water q_i and (ii) $u_i'(t)$ is the marginal willingness to pay for water of agent *i* at the level of consumption *t*. When faced with a two-part tariff for which $T(q_i) = F + \pi q_i$, the calculation of this indicator can be reformulated as:

$$v(q_i) = \int_0^{q_i} u'(t)dt - (F + \pi q_i) = \int_0^{q_i} \left[u'(t) - \pi \right] dt - F \qquad (10.40)$$

Figure 34 shows what the integral measures in the case of a two-part tariff (F, π) for which cost recovery would be incomplete, with $\pi < c + c_e$ and consumption $q_i = q_i^d(F, \pi)$ greater than the first-best value q_i^*. Initially, for each level of consumption $t \in \left[0, q_i^d(F, \pi)\right]$, one computes the difference (vertical distance) between (i) the marginal willingness to pay $u_i'(t)$ for this level of consumption $q_i = t$, that is the maximum sum the household *i* is ready/willing to pay to consume this t-th unit, and (ii) the sum it actually pays, that is the (marginal) price π . In this framework:

- the difference $u'(t) - \pi$ measures (at first approximation) a financial gain made by the agent on this t-th consumption unit;
- the integral $\int_0^{q_i^d(F,\pi)} \left[u'(t) - \pi \right] dt$ by summing the gains made on all the units consumed $t \in \left[0, q_i^d(F, \pi)\right]$ measures the financial gain made by the household, excluding the cost of the subscription, on its drinking water consumption.

Ultimately, the same applies to the net surplus, except that the calculation of the latter takes into account the amount of the subscription F / the lump sum that the household must pay to access to the public drinking water and wastewater services. As apparent, this indicator differs from the household's contribution to the aggregate surplus γ_i . Excluding subscription fee, the consumer surplus for the linear case represented in the set of Figure 34 sets to:

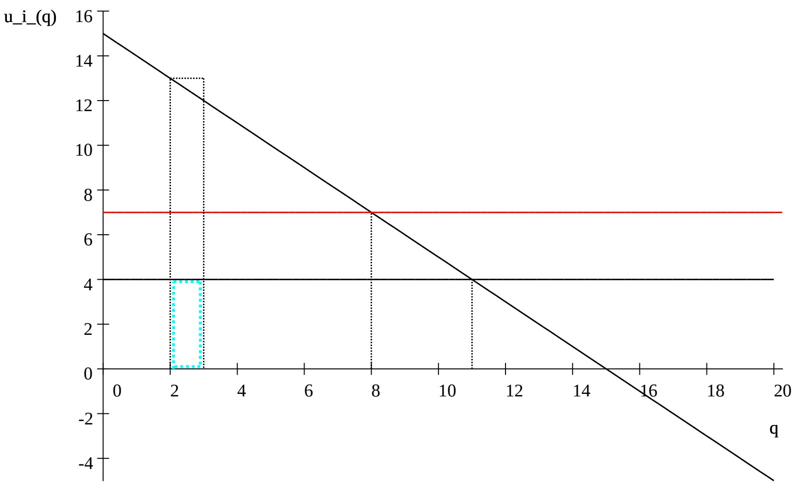
$$\int_0^{q_i} u'(t)dt - \pi q_i = \frac{(\alpha_i - \pi)^2}{2\beta}$$

while the contribution to aggregate surplus sets to:

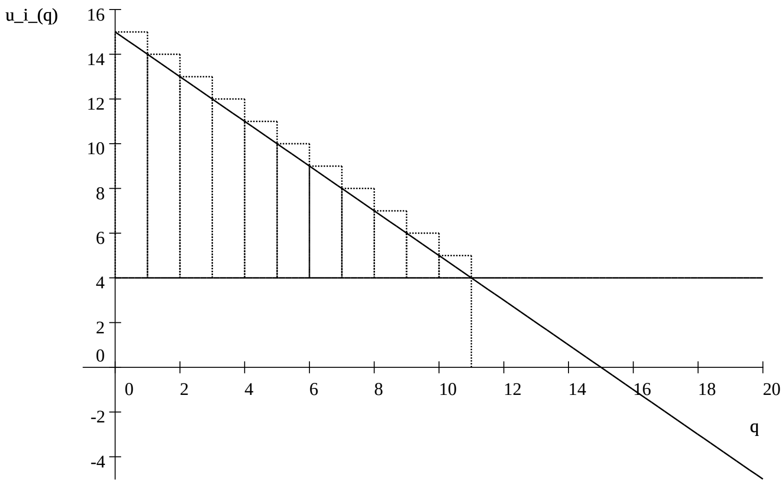
$$\gamma_i = \frac{\left(\alpha_i - (c + c_e)\right)^2}{2\beta} - \frac{(c + c_e - \pi)^2}{2\beta}$$

when the household, facing a two-part tariff (F, π) , sets the level of its consumption to $q_i = q_i^d(F, \pi) = \frac{\alpha_i - \pi}{\beta}$).

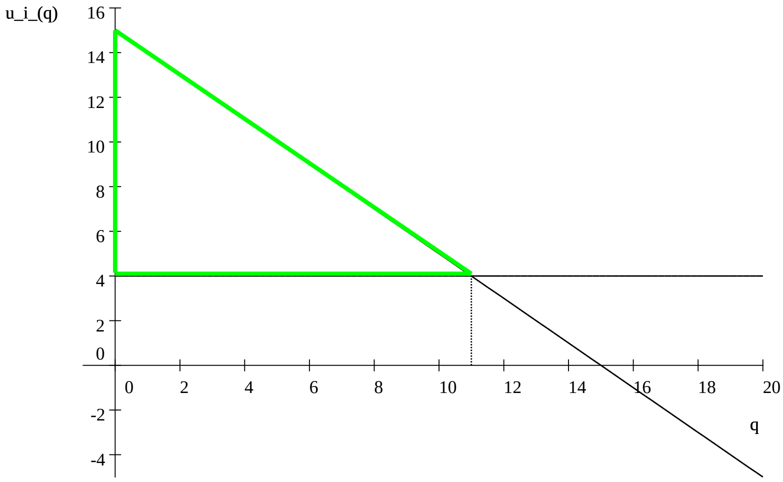
Figure 34 : Computing the consumer surplus



A: calculating a gain - the discrete case



B: calculating overall gain - the discrete case



C: calculation of overall gain - the continuous case ("infinitely divisible")

For the remainder, a technical difficulty is that this sum, given the properties of the approximated water demand function (10.29), takes on an infinite value (*see* the "Technical point" remark on page 141). This leads to compute some variations in surplus by taking the TBSE as the reference point, $\left(F,\pi\right)=\left(\frac{CF}{n},c\right)$, for successively:

- the case of small consumers (located in tranche 1) who are subsidised by the introduction of the IBT.

At a constant $F = CF / n$ subscription level, the latter will then record a gain for sure compared to the TBSE, what is measured by the sum of the (oriented) areas $A_1 > 0$ and. $A_2 > 0$.

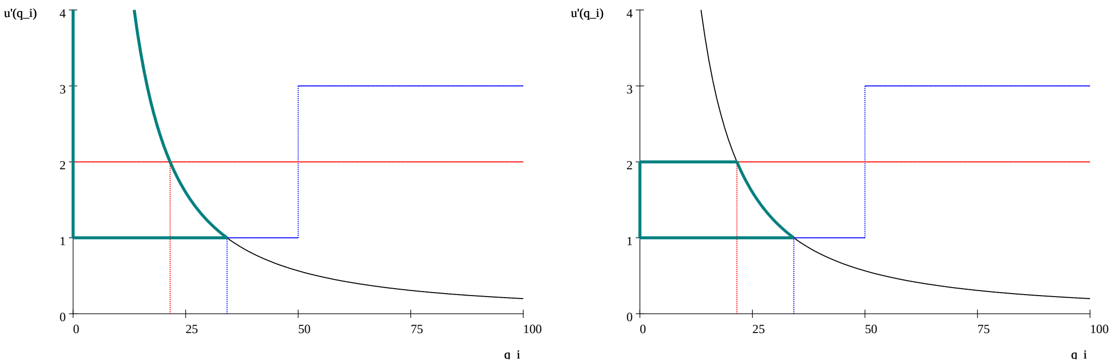


Figure 35 : Variation in consumer surplus (IBT *vs* TBSE) *- the case of the small consumer (gain)*

- the case of large consumers (located in block 2) who are taxed on their block 2 consumption through the introduction of the IBT.

At a constant $F = CF / n$ subscription level, the impact is indeterminate (generally speaking) because of the subsidy for block 1 consumption that also benefits this category of household: *see* the areas oriented $B_1 > 0$, $B_2 < 0$ (curvilinear triangle) and $B_3 < 0$.

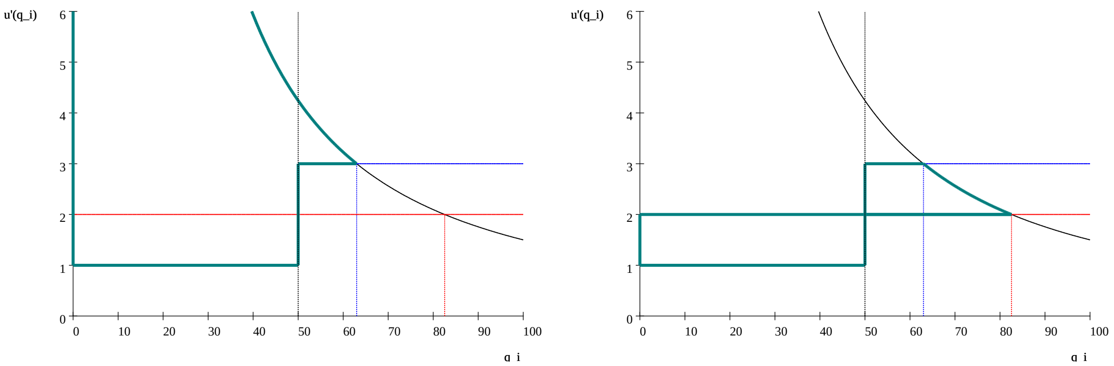


Figure 36 : Variation in consumer surplus (IBT *vs* TBSE) *- the case of the "large" consumer*

showing a gain but a loss is also possible.

- the case of consumers for whom $q^d = k_1$ (corner solution), who are also subsidised on their entire consumption.

When the progressive tariff leads them to reduce their consumption, with $q_i^{\mathrm{IBT}} = k_1 < q_i^{\mathrm{TBSE}}$, the impact (at a constant $F = CF / n$ subscription level) is also indeterminate (generally speaking) because of the subsidy of consumption in block 1: *cf.* the areas oriented $C_1 > 0$ and $C_3 < 0$ for, successively, the case shown in Figure 37 and the case shown in Figure 38 (however, the impact is positive for sure when $q_i^{\mathrm{IBT}} = k_1 > q_i^{\mathrm{TBSE}}$).

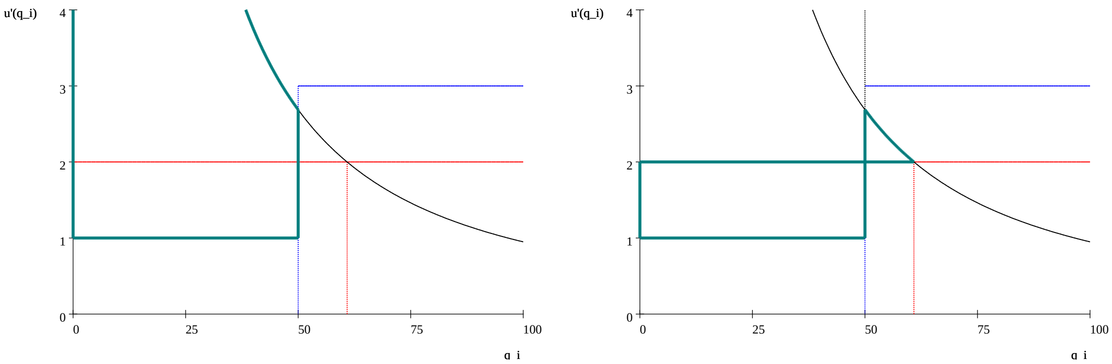


Figure 37 : Variation in consumer surplus (IBT *vs* TBSE) *– corner solution with a gain in fine*

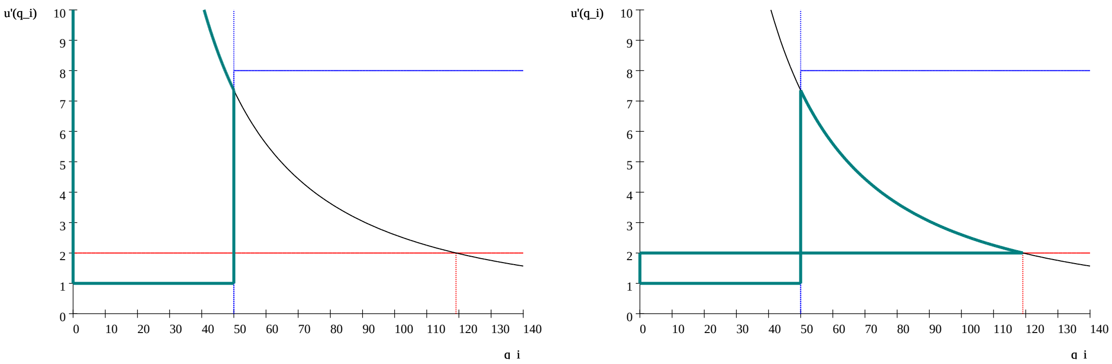


Figure 38 : Variation in consumer surplus (IBT *vs* TBSE) *– corner solution with a loss in fine*

Additional information The change in (net) surplus shown in Figure 35, page 150, can be classically broken down into the sum of two effects:

- the area A_1 reflects the fact that the q_i^{TBSE} units that were previously consumed are now purchased at a lower price (with $\pi_1 < c$), what improves the household's situation;

and at the same time:

- the reduction in price from which the household benefits (everything goes as if) means that it consumes more, and this increase in consumption $q_i^{\mathrm{IBT}} - q_i^{\mathrm{TBSE}} > 0$ generates an additional private gain corresponding to the area of the curvilinear triangle A_2.

At the same time, the impact shown in Figure 36 on page 150 that corresponds to the case of a large consumer is not symmetrical with the previous one because of the non-linear nature of the pricing. Thus:

- The (marginal) price increase faced by the household (everything goes as if) leads it to reduce its consumption, from q_i^{TBSE} to q_i^{IBT}, what reduces its (monetary) well-being by the amount given by the area of the curvilinear triangle B_2.

At the same time, the units it continues to consume were previously purchased:

- at a lower price for the $q_i^{\text{IBT}} - k_1$ units of Block 2, what worsens its situation by an amount given by the area of the rectangle B_3 (equal to $(\pi_2 - c)(q_i^{\text{IBT}} - k_1)$),
- at a higher price for the k_1 units of Block 1, what improves its situation by an amount given by the area of the rectangle B_1 (equal to $(c - \pi_1)k_1$).

The combination of these 3 factors means that the total effect generated by a switch from a TBSE to an IBT is effectively, as a general rule, indeterminate for the category of "large" consumers (Figure 36 show a situation in which the household benefits in fine from a gain but cases in which the total effect is in fine negative can be derived without additional difficulty).

Finally, in the case of the corner solution $q_i^{\text{IBT}} = k_1$ with $q_i^{\text{IBT}} = k_1 < q_i^{\text{TBSE}}$:

- the area of the rectangle C_1, precisely equal to $(c - \pi_1)k_1$, reflects a gain linked to the fact that the k_1 first units that were previously consumed are now purchased at a lower price, what improves the household's situation;

but at the same time:

- the calibration of the pricing system has led (here) the agent to reduce his consumption, from $q_i^{\text{TBSE}} > k_1$ to $q_i^{\text{IBT}} = k_1$, what reduces his (monetary) well-being by the amount given by the area C_2.

As in the previous case, the combination of these two effects means that the total effect is, in the end, indeterminate with, in the end, a gain in the case considered in Figure 37 and a loss for the case considered in Figure 38. ■

Once these elements of analysis provided, the tool computes these losses and gains in net household surplus for each household in the Population file. Moreover:

- the effect of subsidies / "taxation" on the access fee $\frac{CF}{n} - F$,
- the effect linked to the potential direct taxation/subsidy generated by the potential positive/negative operating result (*see* paragraph 9.2 and the IBT-AE)
- the impact of taxation (VAT and excise duties) on water consumption.

are also highlighted. As with variations in the contribution to aggregate surplus, the tool provides basic descriptive statistics on the distribution of these gains and losses borne privately by households.

It should also be kept in mind that these variations in consumer surplus, which measure the monetary gains and losses (of well-being) made/borne by households, differ from the variations in their contribution to aggregate surplus. In particular, the introduction of an IBT may well result in a monetary gain for a consumer in question and a loss for the aggregate surplus (the social welfare), as is the case for instance for the "small" consumers located in block 1 following the introduction of an IBT2 with $\pi_1 < c < \pi_2$.

10.2.3.2 The impact of overconsumption ($\kappa < 1$)

Unlike their effects on the aggregate surplus which are indeterminate (as far as the "taxed" consumption blocks are concerned; see above), the introduction of over-consumption, linked to a poor perception of the tariff, contributes to a definite reduction in the surplus of consumers in block 2 and above, with an additional loss that is precisely given by the area of a curvilinear triangle, such as the one drawn in blue in Figure 39-C, page 154.

One will note that this impact is not limited to the costs of poor management (shown by the area of the dotted red rectangle in the same Figure 39-C).

10.2.3.3 The point reference

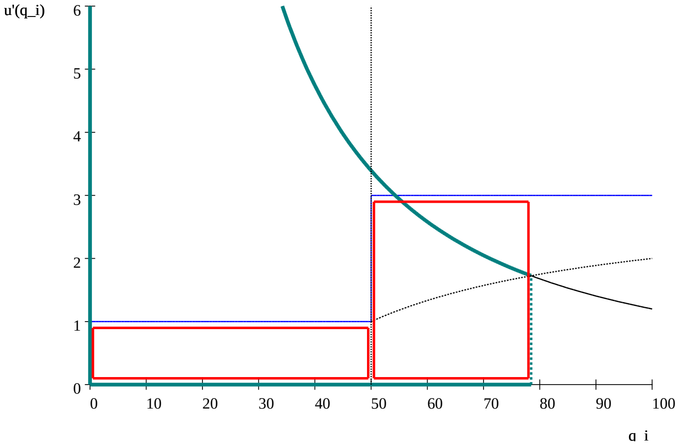
As with contributions to aggregate surplus, the tool measures these variations in consumer surplus (generated by the IBT) by considering an "effective" TBSE in which households connected to the sewerage network face a “EPA” TBSE pricing and those that are not face a “EP” TBSE pricing with regard to the values entered by the user.

10.3 Additional items

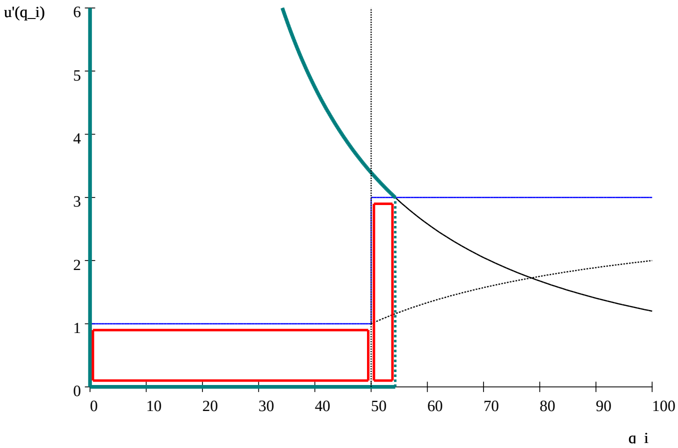
The impact of taxation After presentation of the results for the general population, with impacts calculated for subscription amounts and prices per cubic metre including tax, the calculation of (i) household contributions to the aggregate surplus and (ii) variations in consumer surplus is carried out (in a second stage) excluding charges and VAT, in order to identify the effect of the pricing policy implemented by the Operator, then reproduced for prices including environmental charges but not VAT, so as to identify, by balance, the effect of the ecological tax introduced by the local Water Agency and, finally, for prices including environmental charges and VAT (one then returns to values including VAT) so as to identify, by balance, the effect of the consumption tax mechanism set by the State.

It should be noted that, while the impact of taxation on consumer surpluses is systematically negative, the same is not true for the aggregate surplus, with positive impacts for small consumers (who, compared to the first-order optimum, consume too much) and negative impacts for large consumers (who, compared to the first-order optimum, consume too little), with a total effect that, in the end, is indeterminate.

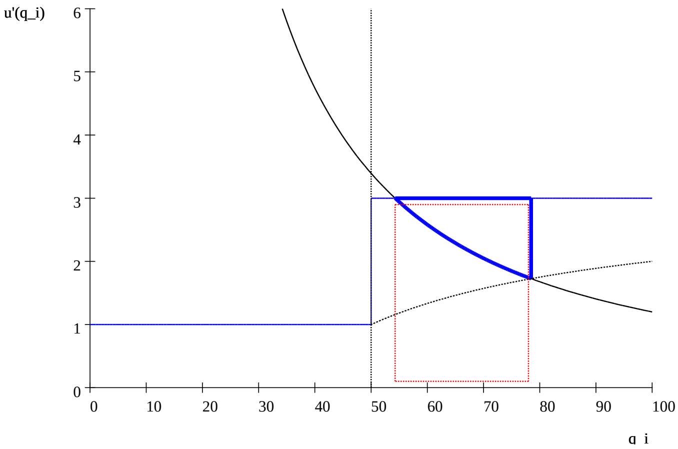
Figure 39 : Consumer surplus - the impact of over-consumption



A: $\kappa = 0$ ("optimum" consumption at stationary equilibrium)



B: $\kappa = 1$ ("optimum" consumption (also) at stationary equilibrium)



C: Loss of consumer surplus (in blue)

Focus and group breakdown As with the previous fields, the tool provides a more detailed (level 2) information by breaking down the subscriber population into 2 groups, and next into 4 groups based on the criteria :

- Group 1 ("The household is not connected to the sewerage network and only pays for drinking water") vs. Group 2 ("The household is connected to the sewerage network and pays the "EPA" tariff"),

- Poor ("The household's standard of living is below the poverty threshold entered by the user") vs. Non Poor ("The household's standard of living is greater than the threshold entered by the user and it is not part of the group of the most deprived households"),

- Poor-G1, Poor-G2, Non Poor-G1 and Non Poor-G2 (obtained by crossing the two criteria described above).

The information in question then refers to focus and group decomposition of suitable indicators (mean, variance, Gini index). Given the specific nature of the problem, this typology is enriched by a third criterion relating to the position of households in relation to the first-best optimum (with a distinction between "small" consumers who consume more than their first-best value, and "large" consumers who consume less than their first-best value).

Recovery of environmental costs Finally, the tool concludes by assessing the recovery of environmental costs, in relation to the values entered by the user, for the general population and two household typologies (G1 vs. G2, on the one hand, G1-Poor / G2-Poor / G1-Non Poor / G2 Non Poor on the other hand) for, respectively, the IBT and the "effective" TBSE.

XI – EVALUATION - QUALITY OF THE FUNDING

11.1 Income statement

The tool first displays the calculation of the operating result with, on the one hand, the calculation and accounting breakdown of fixed costs/variable costs for the EP service, the A service and the overall service:

$$C_{\mathrm{EP}} = C(Q_{\mathrm{EP}}) = CF_{\mathrm{EP}} + c_{\mathrm{EP}} Q_{\mathrm{EP}} \quad (11.1)$$

$$C_{\mathrm{A}} = C(Q_{\mathrm{A}}) = CF_{\mathrm{A}} + c_{\mathrm{A}} Q_{\mathrm{A}} \quad (11.2)$$

$$C_{\mathrm{EPA}} = C(Q_{\mathrm{EP}}, Q_{\mathrm{A}}) = C_{\mathrm{EP}} + C_{\mathrm{A}} = CF + c_{\mathrm{EP}} Q_{\mathrm{EP}} + c_{\mathrm{A}} Q_{\mathrm{A}} \quad (11.3)$$

with $CF = CF_{\mathrm{EP}} + C_{\mathrm{A}}$ and, secondly, the breakdown of company's sales including VAT with the amount of revenue net of tax, the amount of excise duties paid to the Water Agency and the amount of VAT collected for Public Authorities:

$$R_{\mathrm{EP}}^{\mathrm{TTC}} = \sum_{i=1}^{n} T_{\mathrm{EP}}^{\mathrm{TTC}}(q_i^{\mathrm{IBT}}) = n \times F_{\mathrm{EP}} + \sum_{i=1}^{n} \left(\pi_i^{\mathrm{EP}} q_i^{\mathrm{IBT}} - D_i^{\mathrm{EP}} \right) + r_{\mathrm{EP}} Q_{\mathrm{EP}} + \sum_{i=1}^{n} \mathrm{TVA}_i^{\mathrm{EP}} \quad (11.4)$$

$$R_{\mathrm{A}}^{\mathrm{TTC}} = \sum_{i=1}^{n_2} T_{\mathrm{A}}^{\mathrm{TTC}}(q_i^{\mathrm{IBT}}) = n_2 \times F_{\mathrm{A}} + \sum_{i=1}^{n_2} \left(\pi_i^{\mathrm{A}} q_i^{\mathrm{IBT}} - D_i^{\mathrm{A}} \right) + r_{\mathrm{A}} Q_{\mathrm{A}} + \sum_{i=1}^{n_1} \mathrm{TVA}_i^{\mathrm{A}} \quad (11.5)$$

with:

- n_2 the number of household subscribers to the public wastewater service (that is, the size of Group 2),
- π_i^{EP} (mutatis mutandis π_i^{A}) the marginal price (excluding tax) for consumption block *j* in which household *i* is located,
- D_i^{EP} (mutatis mutandis D_i^{A}) the value of Nordin's D (calculated on the basis of prices excluding tax) for this same consumption block *j*.

Noting $T_{\mathrm{EP}}(q_i)$ (mutatis mutandis $T_{\mathrm{A}}(q_i)$) the amount of the bill excluding taxes and charges for the drinking water supply paid by user *i*, the company's operating profit Π is then calculated as:

$$\Pi = \Pi_{\mathrm{EP}} + \Pi_{\mathrm{A}} = \sum_{i=1}^{n} T_{\mathrm{EP}}(q_i^{\mathrm{IBT}}) - C_{\mathrm{EP}}(Q_{\mathrm{EP}}^{\mathrm{IBT}}) + \sum_{i=1}^{n_2} T_{\mathrm{A}}(q_i^{\mathrm{IBT}}) - C_{\mathrm{A}}(Q_{\mathrm{A}}^{\mathrm{IBT}}) \quad (11.6)$$

with:

$$\Pi_{\mathrm{EP}} = n \times \left(F_{\mathrm{EP}} - \frac{CF_{\mathrm{EP}}}{n} \right) + \sum_{i=1}^{n} \left[\left(\pi_i^{\mathrm{EP}} - c_{\mathrm{EP}} \right) q_i^{\mathrm{IBT}} - D_i^{\mathrm{EP}} \right] \quad (11.7)$$

$$\Pi_{\mathrm{A}} = n_1 \times \left(F_{\mathrm{A}} - \frac{CF_{\mathrm{A}}}{n_2} \right) + \sum_{i=1}^{n_2} \left[\left(\pi_i^{\mathrm{A}} - c_{\mathrm{A}} \right) q_i^{\mathrm{IBT}} - D_i^{\mathrm{A}} \right] \quad (11.8)$$

This information is then summarised in an income statement, that also displays some averages per subscriber and per service unit. These figures are expressed on a full-year basis.

Numerical example See Table 35, Table 36, Table 37, Table 38 and Table 39, on next pages, for an numerical example with:

- Table 35, page 158, shows the Operator cost with a Drinking Water line, a Wastewater line and an "Ensemble" line for the consolidated service, also known as the "EP/EPA" service;
- Table 36, page 158, shows Operator revenue with a Drinking Water line and two subheadings, revenue from marketing the service to households not connected to the wastewater service (Group 1) and revenue from marketing the service to households connected to the wastewater service (Group 2), a Wastewater line and a "Ensemble" line for the consolidated "EP/EPA" service;
- Table 37, page 159, gives an account of the operating result with a line relating to the margin generated on the EP (drinking water) service, broken down into a specific subline for Group 1 and a specific subline for Group 2, a line relating to the margin generated on the A (wastewater) service, and a final consolidated line for the general EP / EPA service.

For the case under consideration, these data inform that the operator recorded a (slight) deficit of 179'449 € which represents:

$$\frac{\Pi}{C} = \frac{\Pi_{\mathrm{EP}} + \Pi_{\mathrm{A}}}{C_{\mathrm{EP}} + C_{\mathrm{A}}} = -\frac{-179'449}{22'505'408} = -0.8\%$$

of the total cost of the service (this latter ratio is also calculated and displayed by the tool). As is apparent, this (slight) deficit is due to :

- a loss on the Group 2 Household segment (which pays for drinking water and wastewater services), $\Pi_2 = -1'064'451$,

which is greater (in absolute terms) than :

- the margin generated on the Group 1 Household segment (which is not connected to the sewerage network / only pays for drinking water), $\Pi_1 = 875'372$.

Besides:

- The budget for the wastewater service is virtually balanced, with a profit of $\Pi_{\text{A}} = \Pi_2^{\text{A}} = 5'630$ euros over a financial year,

so that this (slight) deficit for the general service is based on a loss linked to the provision of the public drinking water service for households in group G2 and, given the subsidy for the first consumption blocks (for the IBT tariff which is being tested/assessed here by the user), on an insufficient level of consumption by households in group G2, potentially linked to the higher tariff they face with the EPA service.

As is apparent, in addition to the masses (total cost, aggregate sales, profit) which are broken down into a fixed and a variable component, all of these quantities are also expressed:

- on average per subscriber with, for example, an annual loss to the drinking water service of €3.87 per household, broken down as follows:

$$\frac{\Pi}{n} = \frac{n_1}{n} \times \frac{\Pi_1}{n_1} + \frac{n_2}{n} \times \frac{\Pi_2}{n_2} = 0.471 \times 38.82 + 0.529 \times (-41.92) = 18.30 + (-22.16) \quad (11.9)$$

with €18.30 the average net contribution (in euros) of households in Group 1 and -22.16 the average net contribution (in euros) of households in Group 2 to the profit per subscriber (household) equal here to -3.87 €;

- on average per unit of service with, for example, a margin rate on the EP service that breaks down as follows :

$$\frac{\Pi_{\text{EP}}}{Q} = \frac{Q_1}{Q} \times \frac{\Pi_{\text{EP}}^1}{Q_1} + \frac{Q_2}{Q} \times \frac{\Pi_{\text{EP}}^2}{Q_2} = 0.526 \times 0.24 + 0.474 \times (-0.32) = 0.13 + (-0.15) \quad (11.10)$$

i.e. for 474 units of EP service supplied to Group 2 households, for which a loss of 32 centimes per service unit is achieved (subsidy/taxation on the Access Fee Included), there are 526 units of EP service supplied to Group 1 households, for which a margin of 24 centimes per service unit is achieved (subsidy/taxation on the Access Fee Included);

when these averages are well defined (the drinking water service provided to Group 2 households, which includes a sanitised part unlike Group 1 households, can be considered as a better quality service with EP-G1 and EPA-G2 services which are not homogenous from an economic point of view). The breakdown can then be pushed one step further with :

$$\frac{\Pi_{\text{EP}}}{Q} = \frac{n_1}{n} \times \frac{\bar{q}_1}{\bar{q}} \times \left(\frac{1}{\bar{q}_1} \frac{\Pi_{\text{EP}}^1}{n_1} \right) + \frac{n_2}{n} \times \frac{\bar{q}_2}{\bar{q}} \times \left(\frac{1}{\bar{q}_2} \frac{\Pi_{\text{EP}}^2}{n_2} \right) = 0.471 \times 1.115 \times \frac{38.82}{160.537} + 0.529 \times 0.897 \times \frac{-41.92}{129.159}$$

	n	Q		CF	$\frac{CF}{n}$	CVM	$\frac{CV}{n}$	CV	C	$\frac{C}{n}$	$\frac{C}{Q}$
EP service	47847	6887122		9000000	188.10	0.9	129.55	6198409.81	15198409.81	317.65	2.21
A service	25300	3267494		6000000	237.15	0.4	51.66	1306997.71	7306997.71	288.81	2.24
Total	***	***		15000000		***		7505407.52	22505407.52		***

Table 35 : Operator Cost

	n	Q		nF	F	RVM	$\frac{RV}{n}$	RV	R	$\frac{R}{n}$	$\frac{R}{Q}$
EP service	47847	6887122		3577041.72	74.76	1.66	**239.02**	11436289	15013331	313.78	2.18
G1	22547	3619628		1685613.72	74.76	1.75	281.71	6351723	8037337	356.47	2.22
G2	25300	3267494		1891428	74.76	1.56	200.97	5084566	6975994	275.73	2.13
A service	25300	3267494		1573154	62.18	1.76	226.86	5739474	7312628	289.04	2.24
Total		***		5150196		***		17175763	22325959		***

Table 36 : Operator Revenues

	n	Q		$CF - nF$	$\frac{CF}{n} - F$	$\frac{m_q}{\bar{q}}$	m_q	$n \times m_q$	Π	$\frac{\Pi}{n}$	$\frac{\Pi}{Q}$
EP service	47847	6887122		-5422958.28	-113.34	0.76	109.47	5237879.13	-185079.15	-3.87	-0.03
G1	22547	3619628		-2555467.23	-113.34	0.95	152.16	3430838.71	875371.49	38.82	0.24
G2	25300	3267494		-2867491.05	-113.34	0.55	71.42	1807040.42	-1060450.64	-41.92	-0.32
A service	25300	3267494		-4426846.00	-174.97	1.36	175.20	4432476.22	5630.22	0.22	0.002
Total	***	***		-9849804.28	***	***		9670355.35	**-179448.93**		***

Table 37 : Operating Profit

Identification of G1 EP – G2 EP – G2-A transfers :

$$\frac{\Pi}{n}=\frac{n_1}{n}\times\frac{\Pi_1^{EP}}{n_1}+\frac{n_2}{n}\times\left(\frac{\Pi_2^{EP}}{n_2}+\frac{\Pi_2^{A}}{n_2}\right)=0.471\times 38.82+0.529\times\left(-41.92+0.22\right)=18.30+\left(-22.16\right)+0.12$$

$$\frac{\Pi_{EP}}{Q_{EP}}=\frac{Q_1^{EP}}{Q_{EP}}\times\frac{\Pi_{EP}^1}{Q_1^{EP}}+\frac{Q_2^{EP}}{Q_{EP}}\times\frac{\Pi_{EP}^2}{Q_2^{EP}}=0.526\times 0.24+0.474\times\left(-0.32\right)=0.13+\left(-0.15\right)$$

$$\frac{\Pi_{EP}}{Q_{EP}}=\frac{n_1}{n}\times\frac{\bar{q}_{EP}^1}{\bar{q}_{EP}}\times\left(\frac{1}{\bar{q}_1}\frac{\Pi_{EP}^1}{n_1}\right)+\frac{n_2}{n}\times\frac{\bar{q}_{EP}^2}{\bar{q}_{EP}}\times\left(\frac{1}{\bar{q}_2}\frac{\Pi_{EP}^2}{n_2}\right)=0.471\times 1.115\times\frac{38.82}{160.537}+0.529\times 0.897\times\frac{-41.92}{129.159}$$

$$\frac{\Pi_A}{Q_A}=\frac{1}{\bar{q}_2}\times\frac{\Pi_A^2}{n_2}=\frac{1}{129.159}\times 0.22$$

	n	Q		$n \times tF$	$t \times F$		TVA_q	$n \times \text{TVA}_q$	$n \times \text{TVA}$	TVA	$\frac{\text{TVA}}{Q}$
EP service	47847	6887122		75117.88	1.57		**5.38**	257517.62	332635.4913	6.95	0.05
G1	22547	3619628		35397.89	1.57		6.32	142507.64	177905.5311	7.89	0.05
G2	25300	3267494		39719.99	1.57		4.55	115009.97	154729.9603	6.12	0.05
A service	25300	3267494		157315.4	6.218		23.20	587017.37	744332.77	29.42	0.23
Total	***	***		232433.28	***			844534.99	1076968.26	***	***

Table 38 : State Account

	n	Q		$n \times r_0$	r_0		$r \times \bar{q}$	$n \times r\bar{q}$	$n \times (r_0 + r\bar{q})$	$r_0 + r\bar{q}$	r
EP service	47847	6887122		***	***		17.27	826454.64	***	***	0.12
G1	22547	3619628		***	***		19.26	434355.33	***	***	0.12
G2	25300	3267494		***	***		15.50	392099.31	***	***	0.12
A service	25300	3267494		***	***		5.17	130699.77	***	***	0.04
Total	***	***		***	***			957154.41	***	***	***

Table 39 : Water Agency Account

to assess :

- the number of contributors relative to the number of beneficiaries (in this case, 471 contributors for 529 beneficiaries),
- the ratio of the average bases on which the (net) contributions of Group G1 members and the (net) subsidies of Group G2 members are based (in this case, for 111.5 units of service provided (and taxed on average) to a Group 1 member, 89.7 units of service are provided (and subsidised on average) to a Group 2 member),
- the average tax and subsidy rates, per unit of EP service, that are applied ex post to households in Group 1 and those in Group 2.

These accounting breakdowns with the identification of margins per subscriber and per service unit for the EP-G1, EP-G2 and A-G2 clusters are then also calculated and displayed by the tool.

Finally, these Operator accounts are completed with:

- a government account linked to VAT collection (see Table 38, page 160) ;
- a Water Agency account linked to the collection of excise duties (see Table 39, page 160) ;

whose structures are similar to those of the operator's sales account, with in particular:

- a line relating to the collection of VAT (at a specific rate) for the drinking water service, broken down into a heading for Group 1 and a heading for Group 2, with the calculation of an average amount of tax per capita and per service unit;
- a line relating to the collection of VAT (at a specific rate) for the sanitation service (paid only by households in group 2), with also the calculation of an average amount of tax per capita and per service unit;

and similarly for the collection of excise duties. It should be borne in mind that, unlike environmental charges, the mechanism of VAT (which is an ad valorem tax), applied to prices per cubic metre that increase in step with the consumption, makes that tax per cubic metre is also increasing in step (see the "Remark", page 50). Neglecting the (regressive) effect of the subscription fee (which is also subject to VAT), this non-linear effect means that the average rates per unit of service calculated and displayed by the tool differ from the statutory VAT rates (which are entered by the user).

11.2 Financing structure

11.2.1 Nordin Analysis

In order to facilitate the findings, the tool provides an initial analysis of the IBT treatment of the customer portfolio with a reading “à la Nordin” of the operating profit (REX variable). For these purposes, the net contribution to service funding of household *i*, noted $\Pi_i = T_i - C_i$, is rewritten as:

$$\Pi_i = \hat{\Pi}_{i0} + \hat{m}_i \qquad (11.11)$$

with :

- $\hat{\Pi}_{i0} = F - \frac{CF}{n} - D_i$: the "pseudo" subsidy/taxation on access fee,
- D_i : the Nordin D value for household *i*
- $\hat{m}_i = (\pi_i - c) \times q_i = \mu_i \times q_i$: the "pseudo" subsidy/tax on consumption,
- π_i : the value of the marginal price for the same household *i*,
- $\mu_i = \pi_i - c$: the "pseudo" marginal rate of subsidy/taxation on consumption the Household is facing

with tariff parameters corresponding to their Operator values, excluding taxes and charges. The equation (11.11) states the net margin achieved by the Operator on household *i* breaks down into the sum of 2 terms. The first, $\hat{\Pi}_{i0}$, is a fixed part, positive or negative, that decreases in steps with the level of consumption:

$$\hat{\Pi}_{i0} = \begin{cases} F - \frac{CF}{n} & if \quad 0 \le q \le k_1 \\ F - \frac{CF}{n} - k_1(\pi_2 - \pi_1) & if \quad k_1 < q \le k_2 \\ F - \frac{CF}{n} - [k_1(\pi_2 - \pi_1) + k_2(\pi_3 - \pi_2)] & if \quad k_2 < q \le k_3 \\ \ldots \end{cases} \tag{11.12}$$

The second, $\hat{m}_i = (\pi_i - c) \times q_i = \mu_i q_i$, is a variable part which is negative in the subsidised consumption blocks and positive in the "taxed" blocks with:

$$\hat{m}_i = \mu_i \times q_i = (\pi_i - c) \times q_i = \begin{cases} (\pi_1 - c) q_i & if \quad 0 \le q \le k_1 \\ (\pi_2 - c) q_i & if \quad k_1 < q \le k_2 \\ (\pi_3 - c) q_i & if \quad k_2 < q \le k_3 \\ \ldots \end{cases} \tag{11.13}$$

The point is that, from the operator's point of view, all goes as if the funding mechanism:

- subsidises increasingly the access fees of the Household (assuming that $F < \frac{CF}{n}$, what is the most likely case)

with at the same time:

- prices increasingly all the household water consumption according to positioning in the unit price schedule (block 1, block 2 …).

In other words, from the point of view of its net effects,

- an IBT of the social incentive type can be analysed as a super-progressive pricing system[47] but supplemented by a system of subsidies on the access fee that increases in steps with the level of consumption.

On this basis, the tool displays a supplementary table specifying the average values (per subscriber) of net cash flows, incoming and outgoing, by consumption block:

	$\hat{\Pi}_0$	$f_j = \frac{n_j}{n}$	μ_j	Mean $\hat{m}_j$	Mass $\hat{M}_j$	Mass $\hat{\Pi}_j$
Block 1	$F - \frac{CF}{n}$	f_1	$\pi_1 - c$	$\bar{m}_1 = (\pi_1 - c)\bar{q}_1$	$\hat{M}_1$	$\hat{\Pi}_1$
Block 2	$F - \frac{CF}{n} + D_2$	f_2	$\pi_2 - c$	$\bar{m}_2 = (\pi_2 - c)\bar{q}_2$	$\hat{M}_2$	$\hat{\Pi}_2$
Block 3	$F - \frac{CF}{n} + D_3$	f_3	$\pi_3 - c$	$\bar{m}_3 = (\pi_3 - c)\bar{q}_3$	$\hat{M}_3$	$\hat{\Pi}_3$
⋮						
Overall	…					

Table 40 : Contributions to service funding - reading à la Nordin

with $\bar{q}_1$ the average consumption of household customers located in block 1, $\bar{q}_2$ the average consumption of household customers located in block 2, and so on. This table is produced for, successively, the EP service, the A service and the general (EP / EPA) service. See Table 41, page 164, for a numerical example.

Remarks:

(1) The tables given on page 164 correspond to the Operator treatment and, given the VAT mechanism (which is an ad valorem tax), there is also a State treatment and a consolidated treatment (which is obtained by summing the Operator treatment and the State treatment) named as the Subscriber treatment.

The tables relating to State treatment (VAT) and Subscriber treatment (consolidated) are then displayed in the Economic Efficiency field with, in particular, the calculation of consumer surplus (variations).

(2) For analysis purposes, the tool also produces an Operator Treatment Table, A State Treatment Tabble and a Subscriber Treatment Table for EP-G1 service, EP-G2 service and EPA G2 service, with two potentially interesting comparisons:

- EP G1 vs. EP G2 (EP service breakdown) for which relatively higher EP subsidies and relatively lower contributions to EP service funding are expected for group G2, compared with group G1, given the potentially lower average consumption (all other things being equal);

[47] A super-progressive pricing system combines the payment of a fixed part (subscription fee) and a variable part that is proportional to the volume of water consumed, but with a price per cubic metre that increases in step, so that the household pays all its consumption at the marginal price, that is at the price of the block in which its consumption is located. This system has been applied in Tunisia for a long time, and more recently in Morocco, using a no "step back" clause for households located in block 2 (who still face progressive pricing, with block 1 units sold at the block 1 price).

Table 41 : Nordin reading of the income statement – numerical example

EP Service General population	n_j	f_j (%)	$\hat{\Pi}_{0j}$	$\mu_j = \pi_j - c$	$\bar{q}_j$	$\bar{\hat{m}}_j = \mu_j \bar{q}_j$	$n_j\bar{\Pi}_j$	$n_j\hat{\Pi}_{0j}$	$n_j\bar{\hat{m}}_j$	$n_j\bar{\Pi}_j$
Block 1	814	1.7	-113.34	-0.02	49.04	-1.08	-114.42	-92232.62	-932.06	-93164.68
Bolck 2	14152	29.6	-171.00	0.94	97.35	91.41	-79.59	-2419997.47	1272179.71	-1147817.75
Block 3	30551	63.9	-282.48	1.87	158.75	296.54	14.06	-8630142.77	9107613.95	477471.18
Block 4	2330	4.9	-669.36	3.48	262.71	914.22	244.86	-1559455.23	2137887.33	578432.10
Total	47847	100	**-265.47**	**1.82**	**143.94**	261.60	-3.87	-12701828.09	12516748.94	-185079.15

A service Group G2	n_j	f_j (%)	$\hat{\Pi}_{0j}$	$\mu_j = \pi_j - c$	$\bar{q}_j$	$\bar{\hat{m}}_j = \mu_j \bar{q}_j$	$n_j\bar{\Pi}_j$	$n_j\hat{\Pi}_{0j}$	$n_j\bar{\hat{m}}_j$	$n_j\bar{\Pi}_j$
Block 1	723	2.9	-174.97	0.90	53.00	47.70	-127.27	-126481.31	34480.55	-92000.76
Bolck 2	9879	39.0	-224.17	1.72	92.85	159.71	-64.47	-2214627.10	1577754.65	-636872.46
Block 3	14096	55.7	-234.97	1.81	152.86	276.68	41.70	-3312128.49	3899947.09	587818.61
Block 4	602	2.4	-304.57	2.10	260.99	548.08	243.51	-183469.67	330154.50	146684.83
Total G2	25300	100	**-230.70**	**1.79**	**129.15**	230.92	0.22	-5836706.57	5842336.79	5630.22

EP / EPA service (consolidated)	n_j	f_j (%)	$\hat{\Pi}_{0j}$	$\mu_j = \pi_j - c$	$\bar{q}_j$	$\bar{\hat{m}}_j = \mu_j \bar{q}_j$	$n_j\bar{\Pi}_j$	$n_j\hat{\Pi}_{0j}$	$n_j\bar{\hat{m}}_j$	$n_j\bar{\Pi}_j$
Block 1	814	1.7	-268.77	0.79	52.06	41.23	-227.54	-218713.93	33548.49	-185165.44
Bolck 2	14152	29.6	-327.49	2.10	95.73	201.38	-126.11	-4634624.57	2849934.36	-1784690.21
Block 3	30551	63.9	-390.89	2.67	159.59	425.76	34.87	-11942271.26	13007561.05	1065289.79
Block 4	2330	4.9	-748.11	4.02	263.69	1059.35	311.24	-1742924.90	2468041.83	725116.93
Total	47847	100	**-387.45**	**2.67**	**143.94**	383.70	-3.75	-18538534.66	18359085.74	-179448.93

- EP G1 vs. EPA G2 (breakdown of the general service with a focus on customer segments) for which subsidies and contributions to service funding are expected to be relatively higher for members of the group G2 (with a subsidy-taxation system for wastewater in addition to the one for drinking water) in a context where demand is price inelastic.

This information is then used to create an infographic for a more detailed analysis of how the Operator handles its household subscriber portfolio.

Pen's Parades and Funding Profiles This analysis of the operator treatment of the household customer portfolio, with the application of the IBT which is evaluated/tested by the user, is refined with the production of an infographic whose starting point is the Pen's parade of the customer net margin - Access Fee Included of the general EP / EPA service:

$$\Pi = \left(\Pi_1, \Pi_2, \ldots, \Pi_n\right)$$

with $\Pi_1 \leq \Pi_2 \leq \ldots \leq \Pi_n$ the ranked series of the household customer net margins. This Pen's Parade is the result of the merging of two "Funding Profiles" depending on whether one adopts an approach by Service, EP vs. A, or by customer segment, EP-G1 vs. EPA-G2. In the first approach (by Service), two separate diagrams are shown (along with Pen's parade);

(i) the graph of the "Net margin Acces Fee Included " function for the drinking water service:

$$m_{\mathrm{EP}}\left(q\right) = T_{\mathrm{EP}}\left(q\right) - \left(\frac{CF_{\mathrm{EP}}}{n} + c_{\mathrm{EP}} \times q\right) = F_{\mathrm{EP}} - \frac{CF_{\mathrm{EP}}}{n} - D_{\mathrm{EP}}\left(q\right) + \left(\pi_{\mathrm{EP}}\left(q\right) - c_{\mathrm{EP}}\right) \times q \qquad (11.14)$$

($D_{\mathrm{EP}}\left(q\right)$ the Nordin's D function for the EP service, $\pi_{\mathrm{EP}}\left(q\right)$ the marginal price function for the same EP service) augmented by the histogram of household water consumption for the drinking water service alone;

(ii) the graph of the "Net margin Acces Fee Included " function for the wastewater service:

$$m_{\mathrm{A}}\left(q\right) = T_{\mathrm{A}}\left(q\right) - \left(\frac{CF_{\mathrm{A}}}{n_{\mathrm{A}}} + c_{\mathrm{A}} \times q\right) = F_{\mathrm{A}} - \frac{CF_{\mathrm{A}}}{n_{\mathrm{A}}} - D_{\mathrm{A}}\left(q\right) + \left(\pi_{\mathrm{A}}\left(q\right) - c_{\mathrm{A}}\right) \times q \qquad (11.15)$$

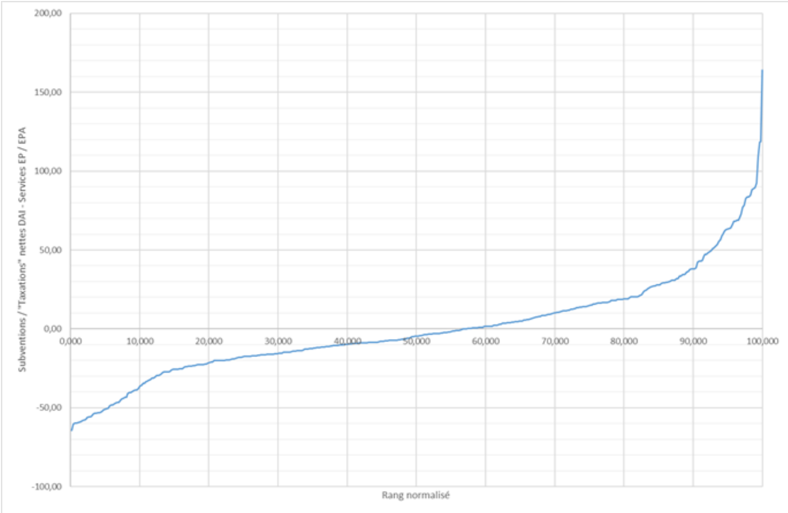
($D_{\mathrm{A}}\left(q\right)$ the Nordin's D function for the A service, $\pi_{\mathrm{A}}\left(q\right)$ the marginal price function for the same A service) augmented by the histogram of household water consumption for the wastewater service alone.

In the second approach (customer segment), two separate diagrams are (still) shown (along with the Pen's parade):

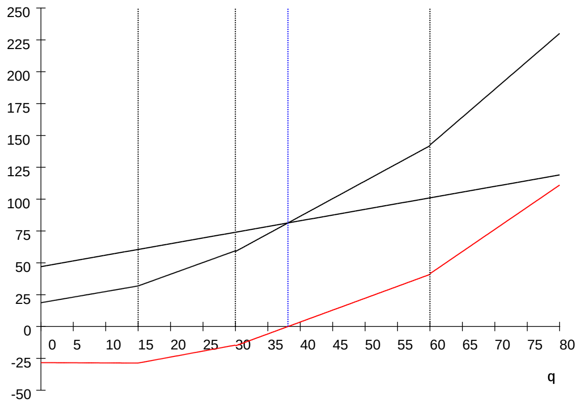
(i) the graph of the "Net margin Acces Fee Included " function for the drinking water service (see equation (11.14)), and the histogram of water consumption for households in group G1 only (not connected to the public sewerage system)

(ii) the graph of the "Net margin Acces Fee Included " function for the water and wastewater service:

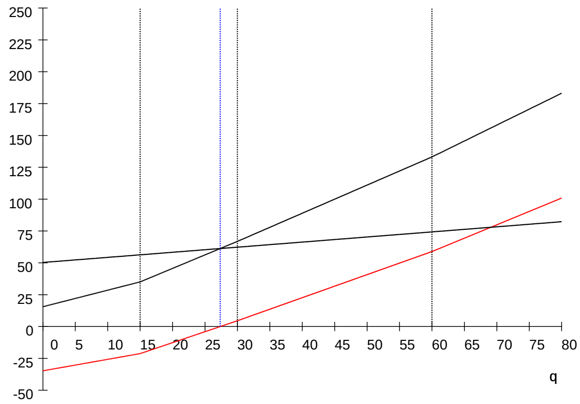
Figure 40 : Breakdown of the Pen's parade of the household net margin by service : EP vs A



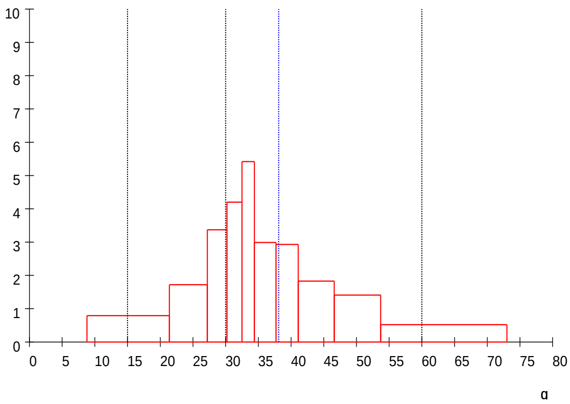
Pen's parade of the household net margin - Access Fee Included, general EP / EPA service



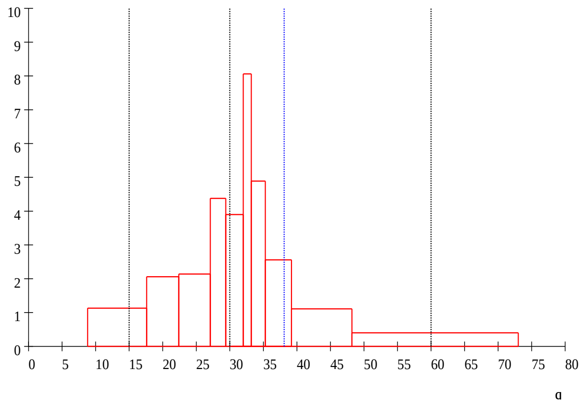
Net margin EP DAI



Net margin A DAI

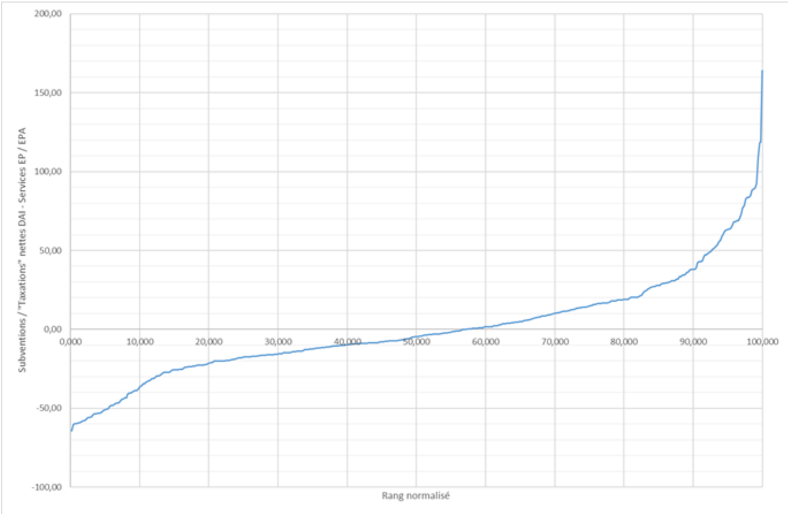


Histogram EP Service

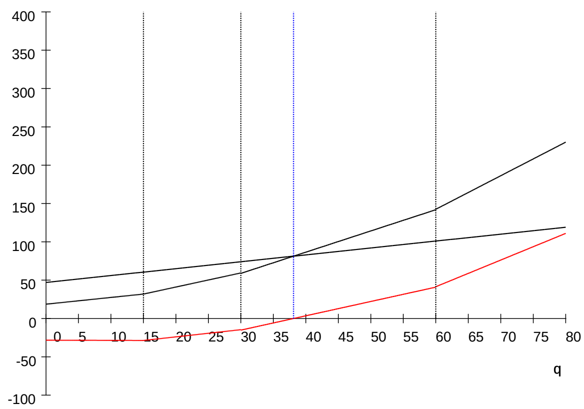


Histogram A Service

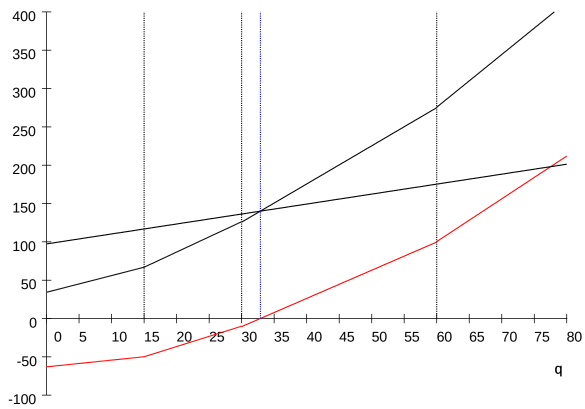
Figure 41 : Breakdown of the Pen's parade of the household net marginby Groups : G1 vs. G2



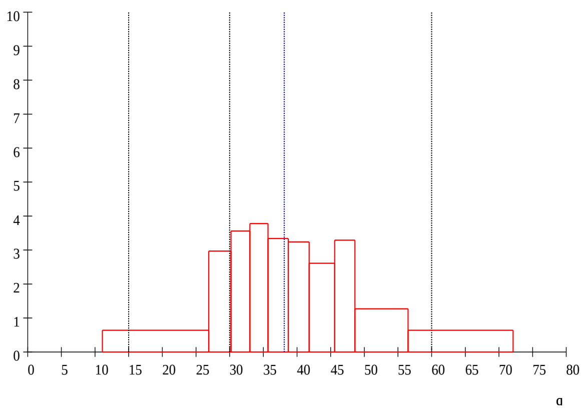
Pen's parade of the household net margin - Access Fee Included, general EP / EPA service



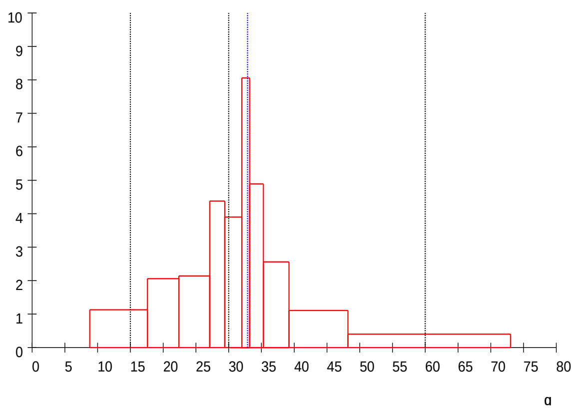
Net margin DAI G1



Net margin DAI G2 (EPA)



Histogram Service G1



Histogram Service G2 (EPA)

$$\begin{aligned} m_{\mathrm{EPA}}(q) &= T_{\mathrm{EP}}(q) + T_{\mathrm{A}}(q) - \left[\frac{CF_{\mathrm{EP}}}{n} + \frac{CF_{\mathrm{A}}}{n_{\mathrm{A}}} + \left(c_{\mathrm{EP}} + c_{\mathrm{A}}\right) \times q\right] \\ &= F_{\mathrm{EP}} + F_{\mathrm{A}} - \left(\frac{CF_{\mathrm{EP}}}{n} + \frac{CF_{\mathrm{A}}}{n_{\mathrm{A}}}\right) - D_{\mathrm{EPA}}(q) + \left(\pi_{\mathrm{EPA}}(q) - c_{\mathrm{EPA}}\right) \times q \end{aligned} \tag{11.16}$$

($D_{\mathrm{EPA}}(q) = D_{\mathrm{EPA}}(q) + D_{\mathrm{A}}(q)$ the Nordin D function for the drinking water and wastewater service, $\pi_{\mathrm{EPA}}(q) = \pi_{\mathrm{EP}}(q) + \pi_{\mathrm{A}}(q)$ the marginal price function for the drinking water and wastewater service) plus the histogram of water consumption for households in group G2 only (connected to the public sewerage system).

As shown, these diagrams display the percentage of households (domestic subscribers) that are net beneficiaries (of the subsidy/taxation system put in place by the IBT) and the percentage of households that are net contributors (to service funding), the amounts of (net) subsidies and (net) "taxes" (and their balancing) that are implemented by the tariff as well as the distribution of these subsidies and "taxes" over the relevant sub-populations, by service (EP vs. A) or by customer segment (Group 1 (EP service only) vs. Group 2 (EPA service)).

11.2.2 Covering fixed costs with variable revenues

11.2.2.1 Décomposition 1 (2 postes)

The aim here is to assess the extent to which fixed costs are financed by variable revenues, and vice versa. For these purposes, it is applied an "economic" breakdown of the operating result (and of the household contributions to service funding) by comparing:

(1) the captive component of household *i*'s revenue $R_i = T(q_i)$, that is:

$$T_{0,i} = T(q_{\grave{a},i}) \tag{11.17}$$

where $q_{0,i} \geq \underline{q}_i$ is the captive part of the consumption of household *i*, to the service cost related to satisfying this captive consumption, that is:

$$C_{0,i} = \frac{CF}{n} + c \times q_{0,i} \tag{11.18}$$

with the calculation of a (potentially negative) net margin:

$$m_{0,i} = T\left(q_{0,i}\right) - \left(\frac{CF}{n} + cq_{0,i}\right) \tag{11.19}$$

(2) the variable part of household *i*'s revenue $R_i = T(q_i)$, that is:

$$T_{q_i - q_{0,i}} = T(q_i) - T(q_{0,i}) \tag{11.20}$$

where $q_i - q_{0,i}$ is the "economic" part of consumption of the household *i*, to the service cost of satisfying this variable consumption, that is:

$$C_{v,i} = c \times (q_i - q_{0,i}) \quad (11.21)$$

with the calculation of a net margin:

$$m_{v,i} = T_{q_i - q_{0,i}} - C_{v,i} = T(q_i) - T(q_{0,i}) - c \times (q_i - q_{0,i}) \quad (11.22)$$

(potentially negative as well).

For example, for a household *i* whose captive consumption $q_{0,i}$ is in block 1 and consumption q_i is in block 2, we have:

	Economic fixed part	Economic variable part	
Revenues	$T_{0,i} = F + \pi_1 q_{0,i}$	$T_{q_i - q_{0,i}} = \pi_1 (q_{0,i} - k_1) + \pi_2 (q_{0,i} - k_1)$	$T_i = T(q_i)$
Service Cost	$C_{0,i} = \frac{CF}{n} + c q_{0,i}$	$C_{v,i} = c(q_i - q_{0,i})$	$C_i = \frac{CF}{n} + c q_i$
Net Margin	$F - \frac{CF}{n} + (\pi_1 - c) q_{0,i}$	$(\pi_1 - c)(q_{0,i} - k_1) + (\pi_2 - c)(q_{0,i} - k_1)$	Π_i

Table 42 : Financing Structure – Household sources

with $\Pi_i = T_i - C_i$ the (possibly negative) margin/profit generated on household *i*. The tool then calculates the sums of these 9 variables (including $T_{0,i}$, $C_{0,i}$, $T_{q_i - q_{0,i}}$ and $C_{v,i}$) for each household in the Population module to display a table of flow that writes:

	Agregate Economic Fixed Part	Agregate Economic Variable part	
Revenues	$R_0 = \sum_{i=1}^{n} T_{0,i}$	$R_v = \sum_{i=1}^{n} T_{q - q_0, i} = \sum_{i=1}^{n} T(q_i) - T(q_{0,i})$	$R = \sum_{i=1}^{n} T(q_i)$
Service Cost	$C_0 = CF + cQ_0$	$C_v = C - C_0 = c \times (Q - Q_0)$	$C = CF + cQ$
Net Margin	$M_0 = R_0 - C_0$	$M_v = R_v - C_v$	$\Pi = R - C$

Table 43 : Aggregate Financing Structure I

where $Q_0 = \sum_{i=1}^{n} q_{0,i}$ is the aggregate captive consumption, and next computes some coverage rates with the ratios:

$$\gamma_0 = \frac{R_0}{C_0} \quad (11.23)$$

$$\gamma_v = \frac{R_v}{C_v} = \frac{R_v}{C - C_0} \quad (11.24)$$

$$\gamma = \frac{R}{C} \quad (11.25)$$

Given the subsidisation of the first consumption blocks, the "taxation" of the higher blocks and the cost recovery constraint, it is expected that:

- the first coverage ratio (linked to the economic fixed charges) takes on a value of less than 1,

For example, a value of 0.4 indicates that the revenue from captive consumption finances 40% of the service cost to the satisfaction of this same captive consumption, with an associated loss of 60-euro cents per euro invested in the production of these units (and for which additional funding has to be found).

- the second ratio (linked to the economic variable charges) takes on a value greater than 1,

For example, a value of 1.2 means that one euro of expenditure invested in the production of non-captive units (here $\frac{1}{c}$ service units) yields 1.20 euro of revenue, of which 20 cents can be allocated to financing the production of captive consumption.

- the overall cost coverage ratio takes a value equal to 1

with the funding of the service which is then well balanced. The tool also displays the following breakdown of the overall cost coverage ratio:

$$\frac{R}{C}=\frac{R_0+R_v}{C_0+C_v}=\frac{C_0}{C_0+C_v}\times\frac{R_0}{C_0}+\frac{C_v}{C_0+C_v}\times\frac{R_v}{C_0+C_v}=\frac{CF+cQ_0}{C}\times\gamma_0+\frac{c\left(Q-Q_0\right)}{C}\times\gamma_v \qquad (11.26)$$

with the first term :

$$\frac{R_0}{C_0+C_v}=\frac{CF+cQ_0}{C}\times\gamma_0$$

which reflects the contribution of the captive part of revenues to cost recovery, and the second term :

$$\frac{R_v}{C_0+C_v}=\frac{c\left(Q-Q_0\right)}{C}\times\gamma_v$$

which reflects the contribution of the variable part of revenue, in the economic sense, to cost recovery. This financing structure can be compared with the one that would emerge, for a given consumption, with the TBSE, i.e. with the shares of captive consumption Q_0 and variable consumption $Q-Q_0$ in the total cost of the service $C(Q)$.

Numerical example See Table 44, page 171, for an illustration relating to the funding of the general service EP / EPA and for which we have:

$$0.992=0.840\times0.623+0.160\times2.934=0.523+0.469$$

This decomposition (which corresponds to the equation (11.26)) states:

- commercial revenues (for the Operator part) cover 99.2% of the cost of the general service (EP / EPA) with (here) a slight deficit of 0.8% of the total (consolidated) cost;
- fixed part of revenues, fixed in the economic sense, that is for the level of service corresponding to the satisfaction of captive consumption Q_0, covers the total cost of the general service $C\left(Q\right)$ up to 52.3%;

Table 44 : Calculating covering ratios – basic breakdown

		Fixed part		Variable part		Total	
n	47847	Average per suscriber	Mass	Average per suscriber	Mass	Average per suscriber	Mass
Revenue		61.53	2943845.31	55.13	2637644.34	116.65	5581489.65
Cost		98.80	4727231.09	18.79	899120.79	117.59	5626351.88
Net margin		-37.27	-1783385.78	36.34	1738523.55	-0.94	-44862.23
Coverage rate			0.623		2.934		0.992
		Cost shares					
			Contributions				Sum
Breakdown		0.840	0.523		0.160	0.469	0.992

$$\gamma = \frac{CF + cQ_0}{C} \times \gamma_0 + \frac{c(Q - Q_0)}{C} \times \gamma_v$$

$\Leftrightarrow \; 0.992 = 0.840 \times 0.623 + 0.160 \times 2.934$

$\Leftrightarrow \; 0.992 = 0.523 + 0.469$

with : $0.992 - 1 = -0.08$

$0.992 - 1 = -0.08$ percentage points,

$0.523 - 0.840 = -31.7$ percentage points,

$0.469 - 0.160 = +30.9$ percentage points,

■

- variable part of revenues, variable in the economic sense, that is generated by the marketing of the variable part of consumption $Q-Q_0$, covers the total cost of the general service $C(Q)$ up to 46.9%.

This financing structure should be compared with the one that emerges with the TBSE, with given IBT consumption, for which:

$$\gamma_{\text{TBSE}} = \frac{CF + cQ_0}{C} \times \gamma_0^{\text{TBSE}} + \frac{c(Q-Q_0)}{C} \times \gamma_v^{\text{TBSE}}$$

$$\Leftrightarrow 1 = 0.840 \times 1 + 0.160 \times 1$$

i.e. i.e. with TBSE (applied to IBT consumptions):

- there is no deficit,
- the fixed part of the revenue, fixed in the economic sense of the term, would cover 84.0% of the total cost of the service,
- the variable part of the revenue, variable in the economic sense of the term, would cover 16.0% of the total cost of the service.

In doing so, the IBT (which is tested/evaluated by the user) generates :

(1) a value for the fixed part of revenues that is below its reference level with (i) a percentage point difference equal to :

$$\Delta\gamma_0 = \frac{C_0}{C_0 + C_v} \times \left(\frac{R_0}{C_0} - 1 \right) = 0.840 \times (0.623 - 1) = 0.523 - 0.840 = -31.7 \text{ points de pourcentage}$$

i.e. compared with TBSE, the fixed share of revenue, fixed in the economic sense of the term, in the financing of the service has been reduced (with the introduction of the IBT) by 31.7 percentage points and (ii) a relative difference which is directly given by :

$$\frac{\Delta\gamma_0}{\gamma_0} = \frac{R_0}{C_0} - 1 = \gamma_0 - 1 = 0.623 - 1 = -0.377 = -37.7\%$$

i.e. compared with the TBSE, fixed share of revenue, fixed in the economic sense of the term, in service funding has been reduced (with the introduction of the IBT) by 37.7%;

(2) the variable part of revenue which is above its reference level (TBSE) with (iii) a percentage point difference which is given by :

$$\Delta\gamma_v = \frac{C_v}{C_0 + C_v} \times \left(\frac{R_v}{C_v} - 1 \right) = 0.160 \times (2.934 - 1) = 0.469 - 0.160 = +30.9 \text{ points de pourcentage}$$

and (iv) a relative deviation (as a % of the IBT-TBSE value) which is directly given by :

$$\frac{\Delta\gamma_v}{\gamma_v} = \frac{R_v}{C_v} - 1 = 2.934 - 1 = 1.934 = +193.4\%$$

i.e. compared with the "IBT-TBSE value", the variable part (in the economic sense of the term) of the revenue is (here) multiplied by almost 3 (compared with what would be obtained if TBSE was applied to these IBT consumptions).

11.2.2.2 Décomposition 2 (4 postes)

This first table on the economic financial structure is then supplemented by a second one showing (1) the basic consumptions (as estimated by the user with the reprocessing of captive components) and (2) the overconsumptions (linked to misperception of the tariff) with a Table 43 that becomes:

	Fixed Part		Variable part		
	Basic	Captive non-basic	$\kappa = 1$	Overconsumption	
Revenues	$\underline{R}$	$R_0 - \underline{R}$	$R_v^{\kappa=1} = \sum_{i=1}^{n} T(q_i^{\kappa=1}) - T(q_{0,i})$	$R_v - R_v^{\kappa=1}$	R
Service Cost	$\underline{C}$	$C_0 - \underline{C}$	$C_v^{\kappa=1} = c \times (Q_{\kappa=1} - Q_0)$	$c \times (Q_{\kappa=\kappa_0} - Q_{\kappa=1})$	C
Net Margin	$\underline{M}$	$M_0 - \underline{M}$	$M_v^{\kappa=1} = R_v^{\kappa=1} - C_v^{\kappa=1}$	$M_v^{\kappa=1} = R_v^{\kappa=1} - C_v^{\kappa=1}$	Π

Table 45 : Aggregate Financing Structure II

with:

- $\underline{R} = \sum_{i=1}^{n} T(\underline{q}_i)$ revenue from the marketing of the basic service, including the collection of subscriptions (access fees),
- $\underline{C} = CF + c\underline{Q}$ the cost to the service of providing the basic service,
- $R_0 - \underline{R} = \sum_{i=1}^{n} T(\underline{q}_i) - \sum_{i=1}^{n} T(q_{0,i})$ revenue from the marketing of non-basic captive consumption, in this case water used for garden maintenance and swimming pool maintenance (excluding user reprocessing),
- $C_0 - \underline{C} = c \times (Q_0 - \underline{Q})$ the cost to the service of producing and distributing these cubic metres for garden and swimming pool maintenance (excluding user reprocessing),
- $R_v^{\kappa=1}$ revenue from the marketing of service units meeting the economic component of demand, excluding over-consumption,
- $C_v^{\kappa=1} = c \times (Q_{\kappa=1} - Q_0)$ the cost of providing the service for the economic component of demand, excluding excess consumption,
- $R_v - R_v^{\kappa=1}$ the contribution of overconsumption to company turnover,

- $c\times\left(Q_{\kappa=\kappa_0}-Q_{\kappa=1}\right)$ the cost to the service generated by the part of the production that corresponds to over-consumption.

The last line of the table shows :

- $\underline{M}=\underline{R}-\underline{C}$: the mass of net margin (likely to be negative) on basic consumption Access Fee Included, that is the (likely) support for the basic service;
- $M_0-\underline{M}=R_0-C_0-\left(\underline{R}-\underline{C}\right)$: the mass of net margins on the captive but not basic part of consumption (the latter is then necessarily Acces Fee Excluded),
- $M_v^{\kappa=1}=R_v^{\kappa=1}-C_v^{\kappa=1}$: the mass of net margin generated on the variable (economic) part of consumption, excluding over-consumption,
- $M_v^{\kappa=1}=R_v^{\kappa=1}-C_v^{\kappa=1}$: the masse of the net margin (likely to be positive) generated on overconsumptions, that is the (presumably) contribution of overconsumption to service funding.

In this context, the decomposition (11.26) becomes:

$$\gamma=\frac{CF+c\underline{Q}}{C(Q)}\times\underline{\gamma}+\frac{c\left(Q_0-\underline{Q}\right)}{C(Q)}\times\gamma_{\text{Fixe\_HB}}+\frac{c\left(Q_{\kappa=1}-Q_0\right)}{C(Q)}\times\gamma_{\text{v}}^{\kappa=1}+\frac{c\left(Q_{\kappa=\kappa_0}-Q_{\kappa=1}\right)}{C(Q)}\times\gamma_{\text{surco}} \quad (11.27)$$

with still $\gamma=R/C$ and:

$$\underline{\gamma}=\frac{\underline{R}}{\underline{C}} \qquad \gamma_{\text{Fixe\_HB}}=\frac{R_0-\underline{R}}{C_0-\underline{C}}=\frac{R_0-\underline{R}}{c\times\left(Q_0-\underline{Q}\right)}$$

$$\gamma_{\text{v}}^{\kappa=1}=\frac{R_{\kappa=1}-R_0}{C\left(Q_{\kappa=1}\right)-C\left(Q_0\right)}=\frac{R_{\kappa=1}-R_0}{c\times\left(Q_{\kappa=1}-Q_0\right)} \qquad \gamma_{\text{surco}}=\frac{R-R_{\kappa=1}}{C(Q)-C\left(Q_{\kappa=1}\right)}=\frac{R-R_{\kappa=1}}{c\times\left(Q-Q_{\kappa=1}\right)}$$

the coverage rates for, respectively, the basic service (with, presumably, $\underline{\gamma}<1$), the captive but non-basic consumption, the economic component of demand excluding over-consumption and the over-consumption (with, presumably, $\gamma_{\text{surco}}>1$). Finally, decomposition (11.27) is reformulated to explicitly show (potential) direct support (normally close to 0):

$$1=\frac{CF+c\underline{Q}}{C(Q)}\times\underline{\gamma}+\frac{c\left(Q_0-\underline{Q}\right)}{C(Q)}\times\gamma_{\text{Fixe\_HB}}+\frac{c\left(Q_{\kappa=1}-Q_0\right)}{C(Q)}\times\gamma_{\text{v}}^{\kappa=1}+\frac{c\left(Q_{\kappa=\kappa_0}-Q_{\kappa=1}\right)}{C(Q)}\times\gamma_{\text{surco}}+\frac{-\Pi}{C(Q)} \quad (11.28)$$

(it is this last breakdown that is displayed in fine by the tool) with $\Pi=R-C$ the operating result of the water company.

Numerical example See Table 46, page 176, for a numerical illustration concerning the funding of the general service EP / EPA, and for which:

$$1=\frac{CF+c\underline{Q}}{C(Q)}\times\underline{\gamma}+\frac{c(Q_0-\underline{Q})}{C(Q)}\times\gamma_{\text{Fixe\_HB}}+\frac{c(Q_{\kappa=1}-Q_0)}{C(Q)}\times\gamma_{\text{v}}^{\kappa=1}+\frac{c(Q_{\kappa=\kappa_0}-Q_{\kappa=1})}{C(Q)}\times\gamma_{\text{surco}}+\frac{-\Pi}{C(Q)}$$
$$=\frac{81.2}{100}\times 0.563+\frac{2.8}{100}\times 2.361+\frac{13.9}{100}\times 2.856+\frac{2.0}{100}\times 3.446+0.008$$
$$=0.457+0.066+0.387+0.072+0.008$$

(11.29)

The interpretation is similar to decomposition 1 (with 2 items) except that the contributions of basic service (which includes the access fee) and overconsumptions to service funding are highlighted. In particular, the equation (11.29) shows:

- commercial revenues (for the Operator part) cover the cost of the (general) service up to $1-0.8\%=99.2\%$ (*cf*. also the value of the coverage rate $\gamma=R/C$ displayed in Table 46) with (here) a slight deficit of 0.8% of the total (consolidated) cost;
- marketing of the basic service covers 45.7% of the cost of the (general) service, with a self-financing rate of $\underline{\gamma}=0.563=56.3\%$,

i.e. the marketing of the basic service funds 56.3% of the Operator costs linked to the provision of the basic service, the latter including the collection of an access fee and a (non-linear) tariff for basic cubic metres. These values should be compared with those obtained for the TBSE (and, where applicable, a service maintained at its IBT level), that is:

$$\left(\frac{\underline{R}}{C}\right)_{\text{IBT\_TBSE}}=\frac{\underline{C}}{C}=0.812$$

$$\underline{\gamma}_{\text{IBT\_TBSE}}=\frac{\underline{R}_{\text{IBT\_TBSE}}}{\underline{C}}=\frac{\sum_{i=1}^{n}\left(F_{\text{TBSE}}+\pi_{\text{TBSE}}\times\underline{q}_i\right)}{CF+c\underline{Q}}=\frac{n\times F_{\text{TBSE}}+\pi_{\text{TBSE}}\sum_{i=1}^{n}\underline{q}_i}{CF+c\underline{Q}}=\frac{n\times\frac{CF}{n}+c\times\underline{Q}}{CF+c\underline{Q}}=1$$

This gives (i) a difference in percentage points equal to :

$$\frac{\underline{R}}{C}-\left(\frac{\underline{R}}{C}\right)_{\text{TBSE}}=\frac{\underline{C}}{C}\times\frac{\underline{R}}{\underline{C}}-\frac{\underline{C}}{C}=\frac{\underline{C}}{C}\times\left(\frac{\underline{R}}{\underline{C}}-1\right)=0.812\times 0.563-0.812=0.812\times(0.563-1)$$
$$=\frac{\underline{C}}{C}\times\underline{\gamma}-\frac{\underline{C}}{C}=\frac{\underline{C}}{C}\times(\underline{\gamma}-1)=0.457-0.812=0.812\times(-0.437)$$
$$=-0.354=-35.4\text{ points de pourcentage}$$

i.e. ceteris paribus (with, in particular, a given level of IBT consumption), the implementation of the IBT reduces the share of revenue linked to the marketing of basic service by 35.5 percentage points with (ii) a relative difference (as a percentage of the IBT-TBSE value) equal to:

Table 46 : Calculation of coverage rates -- breakdown n°2 for General Service "EP / EPA"

	Fixed Part				Variable part					
n	Mean	Mass	Mean	Mass	Mean	Mass	Mean	Mass	Mean	Mass
47847	Basic Consumption		Non basic Captive Consumption		PP Economic Part		Overconsumption			
Revenue	53.76	2572411.44	7.76	371433.87	46.63	2231219.78	8.49	406424.55	116.65	5581489.65
Cost	95.51	4569915.19	3.29	157315.90	16.33	781164.67	2.47	117956.11	117.59	5626351.88
Net margin	-41.75	-1997503.75	4.48	214117.97	30.31	1450055.11	6.03	288468.44	-0.94	-44862.23
Coverage rate		0.563		2.361		2.856		3.446		0.992
	Cost shares									
		Contribution								Sum
Breakdown	0.812	**0.457**	0.028	**0.066**	0.139	**0.397**	0.02	**0.072**		0.992

Identification and Quantification of funding sources for General Service:

$$1=\frac{CF+c\underline{Q}}{C(Q)}\times\underline{\gamma}+\frac{c\left(Q_0-\underline{Q}\right)}{C(Q)}\times\gamma_{\text{Fixe\_HB}}+\frac{c\left(Q_{\kappa=1}-Q_0\right)}{C(Q)}\times\gamma_{\text{v}}^{\kappa=1}+\frac{c\left(Q_{\kappa=\kappa_0}-Q_{\kappa=1}\right)}{C(Q)}\times\gamma_{\text{surco}}+\frac{-\Pi}{C(Q)}$$

$$=\frac{81.2}{100}\times0.563+\frac{2.8}{100}\times2.361+\frac{13.9}{100}\times2.856+\frac{2.0}{100}\times3.446+0.008$$

$$=0.457+0.066+0.387+0.072+0.008$$

■

$$\frac{0.812\times(0.563-1)}{0.812}=0.563-1=-0.437=-43.7\%=\underline{\gamma}-1$$

Examination of other contributions :

$$\frac{R_0-\underline{R}}{C}=0.066=\frac{C-\underline{C}}{C}\times\frac{R_0-\underline{R}}{C-\underline{C}}=0.028\times 2.361$$

$$\frac{R_v^{\kappa=1}}{C}=0.397=\frac{C_v^{\kappa=1}}{C}\times\frac{R_v^{\kappa=1}}{C_v^{\kappa=1}}=0.139\times 2.856$$

$$\frac{R_v^{\text{surco}}}{C}=0.072=\frac{C_v^{\kappa=\kappa_0}-C_v^{\kappa=1}}{C}\times\frac{R_v^{\kappa=1}}{c\times\left(Q_{\kappa=\kappa_0}-Q_{\kappa=1}\right)}=0.02\times 3.446$$

shows:

- captive but non-basic consumption, in this case garden and swimming pool maintenance, contributes 6.6% to the recovery of (general) service costs, with a multiplier of 2.361, i.e. 1 euro spent on satisfying captive but non basic consumption generates an average of 2,361 euros in revenue to fund the (general) service;
- economic part of consumption (excluding over-consumption) contributes 39.7% of the recovery of (general) service costs, with a multiplier of 2.856;
- over-consumption (due to tariff misperception) contributes 7.2% to the recovery of (general) service costs, with a multiplier of 3.446.

The values of these contributions are then $0.066-0.028=3.8$ percentage points, $0.397-0.139=25.8$ percentage points and $0.072-0.02=5.2$ percentage points higher than the TBSE reference values, i.e.:

- ceteris paribus (with, in particular, given IBT consumption), the introduction of IBT increased the weight of these consumptions in the funding of general service by 3.8 percentage points, 25.8 percentage points and 5.2 percentage points respectively.

Lastly, the deficit equal (in this case) to 0.8% of the cost of the service constitutes direct support (funded by other sources of funding, necessarily).

Finally, this information (production of Table 46 + calculation of coverage rates + Disaggregation of the overall coverage rate into its various components + display of direct support) is also shown:

- for the EP service and the A service (G2) with a focus on Services (Activities),
- for the G1 group (EP) and the G2 group (EPA) with a focus on "Customer Segments",
- for groups G1 and G2 for the EP service only

so as to cover all the a priori comparisons that the user would like to be able to perform. On this point, see Table 47 page 178, Table 48 page 179, Table 49 page 180, Table 50 page 181 and Table 51 page 182.

Table 47 : Calculation of coverage rates -- breakdown n°2 for "EP" Service

	Fixed Part				Variable part					
n	Mean	Mass	Mean	Mass	Mean	Mass	Mean	Mass	Mean	Mass
47847	Basic Consumption		Non basic Captive Consumption		PP Economic Part		Overconsumption			
Revenue	33.92	1 622 923.13	4.97	237 563.98	33.15	1 586 092.12	6.41	306 753.44	78.44	3 753 332.67
Cost	60.85	2 911 385.02	2.72	130 182.26	13.76	658 217.29	2.09	99 817.88	79.41	3 799 602.45
Net margin	-26.93	-1 288 461.89	2.24	107 381.71	19.39	927 874.83	4.32	206 935.56	-0.97	-46 269.79
Coverage rate		0.557		1.825		2.410		3.073		0.988
	Cost shares									
		Contribution								Sum
Breakdown	0.766	**0.427**	0.034	**0.063**	0.173	**0.417**	0.03	**0.081**		0.988

Identification and quantification of funding sources for EP Service:

$$1=\frac{CF_{\mathrm{EP}}+c_{\mathrm{EP}}\underline{Q}}{C_{\mathrm{EP}}(Q)}\times\underline{\gamma}^{\mathrm{EP}}+\frac{c_{\mathrm{EP}}\left(Q_0-\underline{Q}\right)}{C_{\mathrm{EP}}(Q)}\times\gamma^{\mathrm{EP}}_{\mathrm{Fixe\_HB}}+\frac{c_{\mathrm{EP}}\left(Q_{\kappa=1}-Q_0\right)}{C_{\mathrm{EP}}(Q)}\times\gamma_{\mathrm{v}}^{\kappa=1,\mathrm{EP}}+\frac{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0}-Q_{\kappa=1}\right)}{C_{\mathrm{EP}}(Q)}\times\gamma^{\mathrm{EP}}_{\mathrm{surco}}+\frac{-\Pi_{\mathrm{EP}}}{C_{\mathrm{EP}}(Q)}$$

$$=\frac{76.6}{100}\times 0.557+\frac{3.4}{100}\times 1.825+\frac{17.3}{100}\times 2.410+\frac{3.0}{100}\times 3.073+0.012$$

$$=0.427+0.063+0.417+0.081+0.012$$

■

Table 48 : Calculation of coverage rates -- breakdown n°2 for "A" Service

	Fixed Part				Variable part					
Abonnés	Mean	Mass	Mean	Mass	Mean	Mass	Mean	Mass	Mean	Mass
25300	Basic Consumption		Non basic Captive Consumption		PP Economic Part		Overconsumption			
Revenue	37.53	949 488.31	5.29	133 869.89	25.50	645 127.66	3.94	99 671.11	72.26	1 828 156.98
Cost	65.55	1 658 530.17	1.07	27 133.64	4.86	122 947.38	0.72	18 138.23	72.20	1 826 749.43
Net margin	-28.03	-709 041.86	4.22	106 736.25	20.64	522 180.28	3.22	81 532.88	0.06	1 407.56
Coverage rate		0.572		4.934		5.247		5.495		1.001
	Cost shares									
		Contribution								Somme
Breakdown	0.908	**0.520**	0.015	**0.073**	0.067	**0.353**	0.01	**0.055**		1.001

Identification and quantification of funding sources for A Service (G2) :

$$1=\frac{CF_{A}+c_{A}\underline{Q}_{A}}{C_{A}\left(Q_{A}\right)}\times\underline{\gamma}^{A}+\frac{c_{A}\left(Q_{0,A}-\underline{Q}_{A}\right)}{C_{A}\left(Q_{A}\right)}\times\gamma_{\text{Fixe\_HB}}^{A}+\frac{c_{A}\left(Q_{\kappa=1,A}-Q_{0,A}\right)}{C_{A}\left(Q_{A}\right)}\times\gamma_{v}^{\kappa=1,A}+\frac{c_{A}\left(Q_{\kappa=\kappa_{0},A}-Q_{\kappa=1,A}\right)}{C_{A}\left(Q_{A}\right)}\times\gamma_{\text{surco}}^{A}+\frac{-\Pi_{A}}{C_{A}\left(Q_{A}\right)}$$

$$=\frac{90.8}{100}\times 0.572+\frac{1.5}{100}\times 4.934+\frac{6.7}{100}\times 5.247+\frac{1.0}{100}\times 5.495-0.001$$

$$=0.520+0.073+0.353+0.055-0.001$$

■

Table 49 : Calculation of coverage rates -- breakdown n°2 for "EP"-G1 Service

	Fixed Part				Variable part					
Abonnés	Mean	Mass	Mean	Mass	Mean	Mass	Mean	Mass	Mean	Mass
22547	Basic Consumption		Non basic Captive Consumption		PP Economic Part		Overconsumption			
Revenue	33.44	754 069.64	5.45	122 814.73	41.80	942 549.49	8.42	189 900.30	89.12	2 009 334.16
Cost	60.54	1 364 962.36	3.07	69 131.58	16.92	381 585.69	2.62	59 006.85	83.15	1 874 686.48
Net margin	-27.09	-610 892.72	2.38	53 683.15	24.88	560 963.80	5.81	130 893.45	5.97	134 647.68
Coverage rate		0.552		1.777		2.470		3.218		1.072
	Cost shares									
		Contribution								Sum
Breakdown	0.728	**0.402**	0.037	**0.066**	0.204	**0.503**	0.03	**0.101**		1.072

Identification and quantification of funding sources for G1 (EP Service):

$$1=\frac{CF_{\mathrm{EP}}+c_{\mathrm{EP}}\underline{Q}_1}{C_{\mathrm{EP}}\left(Q_1\right)}\times\gamma_{\underline{1}}^{\mathrm{EP}}+\frac{c_{\mathrm{EP}}\left(Q_{0,1}-\underline{Q}_1\right)}{C_{\mathrm{EP}}\left(Q_1\right)}\times\gamma_{\mathrm{Fixe\_HB},1}^{\mathrm{EP}}+\frac{c_{\mathrm{EP}}\left(Q_{\kappa=1,1}-Q_{0,1}\right)}{C_{\mathrm{EP}}\left(Q_1\right)}\times\gamma_{\mathrm{v},1}^{\kappa=1,\mathrm{EP}}+\frac{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0,1}-Q_{\kappa=1,1}\right)}{C_{\mathrm{EP}}\left(Q_1\right)}\times\gamma_{\mathrm{surco},1}^{\mathrm{EP}}+\frac{-\Pi_{\mathrm{EP},1}}{C_{\mathrm{EP}}\left(Q_1\right)}$$

$$=\frac{72.8}{100}\times0.552+\frac{3.7}{100}\times1.777+\frac{20.4}{100}\times2.470+\frac{3.0}{100}\times3.218-0.072$$

$$=0.402+0.066+0.503+0.101-0.072$$

■

Table 50 : Calculation of coverage rates -- breakdown n°2 for "EP"-G2 service

	Fixed Part				Variable part					
Abonnés	Mean	Mass	Mean	Mass	Mean	Mass	Mean	Mass	Mean	Mass
25300	Basic Consumption		Non basic Captive Consumption		PP Economic Part		Overconsumption			
Revenue	34.34	868 853.48	4.54	114 749.25	25.44	643 542.64	4.62	116 853.14	68.93	1 743 998.51
Cost	61.12	1 546 422.65	2.41	61 050.69	10.93	276 631.61	1.61	40 811.03	76.08	1 924 915.97
Net margin	-26.78	-677 569.17	2.12	53 698.56	14.50	366 911.03	3.01	76 042.12	-7.15	-180 917.47
Coverage rate		0.562		1.880		2.326		2.863		0.906
	Cost shares									
		Contribution								Sum
Breakdown	0.803	**0.451**	0.032	**0.060**	0.144	**0.334**	0.02	**0.061**		0.906

Identification and quantification of funding sources for G2-EP Service:

$$1=\frac{CF_{\text{EP}}+c_{\text{EP}}\underline{Q}_2}{C_{\text{EP}}\left(Q_2\right)}\times\underline{\gamma}_2^{\text{EP}}+\frac{c_{\text{EP}}\left(Q_{0,2}-\underline{Q}_2\right)}{C_{\text{EP}}\left(Q_2\right)}\times\gamma_{\text{Fixe\_HB},2}^{\text{EP}}+\frac{c_{\text{EP}}\left(Q_{\kappa=1,2}-Q_{0,2}\right)}{C_{\text{EP}}\left(Q_2\right)}\times\gamma_{\text{v},2}^{\kappa=1,\text{EP}}+\frac{c_{\text{EP}}\left(Q_{\kappa=\kappa_0,2}-Q_{\kappa=1,2}\right)}{C_{\text{EP}}\left(Q_2\right)}\times\gamma_{\text{surco},2}^{\text{EP}}+\frac{-\Pi_{\text{EP},2}}{C_{\text{EP}}\left(Q_2\right)}$$

$$=\frac{80.3}{100}\times0.562+\frac{3.2}{100}\times1.880+\frac{14.4}{100}\times2.326+\frac{2.0}{100}\times2.863+0.094$$

$$=0.451+0.060+0.334+0.061+0.094$$

■

Table 51 : Calculation of coverage rates -- breakdown n°2 for "EPA"-G2 service

	Fixed Part				Variable part					
Abonnés	Mean	Mass	Mean	Mass	Mean	Mass	Mean	Mass	Mean	Mass
25300	Basic Consumption		Non basic Captive Consumption		PP Economic Part		Overconsumption			
Revenue	71.87	1 818 341.80	9.83	248 619.14	50.94	1 288 670.30	8.56	216 524.25	158.43	3 572 155.49
Cost	126.68	3 204 952.83	3.49	88 184.33	15.79	399 578.99	2.33	58 949.26	148.29	3 751 665.40
Net margin	-54.81	-1 386 611.03	6.34	160 434.81	35.14	889 091.31	6.23	157 574.99	-7.10	-179 509.91
Coverage rate		0.567		2.819		3.225		3.673		0.952
	Cost shares									
		Contribution								Sum
Breakdown	0.854	**0.485**	0.024	**0.066**	0.107	**0.343**	0.02	**0.058**		0.952

Identification and quantification of funding sources for "EPA" Service (G2) :

$$
\begin{aligned}
1 &= \frac{CF_{\mathrm{EPA}} + c_{\mathrm{EPA}}\underline{Q}_2}{C_{\mathrm{EPA}}\left(Q_2\right)} \times \gamma_{\underline{2}}^{\mathrm{EPA}} + \frac{c_{\mathrm{EPA}}\left(Q_{0,2} - \underline{Q}_2\right)}{C_{\mathrm{EPA}}\left(Q_2\right)} \times \gamma_{\mathrm{Fixe\_HB},2}^{\mathrm{EPA}} + \frac{c_{\mathrm{EPA}}\left(Q_{\kappa=1,2} - Q_{0,2}\right)}{C_{\mathrm{EPA}}\left(Q_2\right)} \times \gamma_{\mathrm{v},2}^{\kappa=1,\mathrm{EPA}} + \frac{c_{\mathrm{EPA}}\left(Q_{\kappa=\kappa_0,2} - Q_{\kappa=1,2}\right)}{C_{\mathrm{EPA}}\left(Q_2\right)} \times \gamma_{\mathrm{surco},2}^{\mathrm{EPA}} + \frac{-\Pi_{\mathrm{EPA},2}}{C_{\mathrm{EPA}}\left(Q_2\right)} \\
&= \frac{85.4}{100} \times 0.567 + \frac{2.4}{100} \times 2.819 + \frac{10.7}{100} \times 3.225 + \frac{2.0}{100} \times 3.673 + 0.048 \\
&= 0.485 + 0.066 + 0.343 + 0.058 + 0.048
\end{aligned}
$$

■

11.2.2.3 Additional desegregations (Groups, Services and Groups and Services)

The breakdown of the coverage rate for the General Service EP / EPA is refined with the calculation of the contributions linked to the (3) distinctions Groups (G1 vs. G2), Services (EP vs. A) and Groups and Services (EP-G1, EP-G2 and A-G2).

A) Additional disaggregation Returning to equation (11.26):

$$\gamma = \frac{CF + cQ_0}{C} \times \gamma_0 + \frac{c\left(Q - Q_0\right)}{C} \times \gamma_v$$

with:

$$\gamma = \frac{R}{C},\ \gamma_0 = \frac{R_0}{C_0},\ \gamma_v = \frac{R_v}{C_v}$$

the coverage rates of the general EP/EPA service, the captive component and the economic component of household water consumption, this decomposition is next refined by introducing (1) the dimension Groups G1 vs. G2, (2) the distinction Services EP vs. A, and (3) the distinction Groups and Services EP-G1, EP-G2 and A-G2.

1) G1 vs. G2 decomposition For the G1 vs. G2 decomposition, we have:

$$\gamma = \frac{R}{C} = \frac{R_1 + R_2}{C_1 + C_2} = \frac{C_1}{C_1 + C_2} \times \frac{R_1}{C_1} + \frac{C_2}{C_1 + C_2} \times \frac{R_2}{C_2} = \frac{C_1}{C_1 + C_2} \times \gamma_1 + \frac{C_2}{C_1 + C_2} \times \gamma_2 \tag{11.30}$$

with:

- R_1 the Operator revenue (excluding taxes and environmental fees) generated by marketing the EP service to group G1 only (sub-population of subscribers not connected to the collective sanitation network),

- R_2 the Operator revenue (excluding taxes and environmental fees) generated by marketing the EPA service to the G2 group only (sub-population of subscribers connected to the collective sanitation network),

- $C_1 = n_1 \times \frac{CF_{\text{EP}}}{n} + c_{\text{EP}} \times Q_1$ the cost to the service of providing the EP service for group G1,

- $C_2 = n_2 \times \left(\frac{CF_{\text{EP}}}{n} + \frac{CF_{\text{A}}}{n_2}\right) + \left(c_{\text{EP}} + c_{\text{A}}\right) \times Q_2$ the cost to the service of providing the EPA service for group G2

and:

$$\gamma_1 = \frac{R_1}{C_1} = \frac{R_0^1 + R_v^1}{C_0^1 + C_v^1} = \frac{C_0^1}{C_0^1 + C_v^1} \times \frac{R_0^1}{C_0^1} + \frac{C_v^1}{C_0^1 + C_v^1} \times \frac{R_v^1}{C_v^1} = \frac{n_1 \frac{CF_{\text{EP}}}{n} + cQ_0^1}{n_1 \frac{CF_{\text{EP}}}{n} + cQ_1} \times \gamma_0^1 + \frac{c\left(Q_1 - Q_0^1\right)}{n_1 \frac{CF}{n} + cQ_1} \times \gamma_v^1$$

Table 52 : Decomposition by Groups, by Factors, and by Groups and Factors of equation (11.26) *– Summary tables*

Cost share matrix (Weights)									
	EP			A			EP / EPA (consolidated)		
	Fixed part	Variable part	**Aggregate**	Fixed part	Variable part	Aggregate	Fixed part	Variable part	**Aggregate**
Group 1	0.255	0.078	**0.333**	***	***	***	0.255	0.078	**0.333**
Group 2	0.286	0.056	**0.342**	0.300	0.025	0.325	0.585	0.081	**0.667**
All	0.541	0.135	**0.675**	0.300	0.025	0.325	0.840	0.160	**1.000**

Coverage rate matrix									
	EP			A			EP / EPA (consolidated)		
	Fixed part	Variable part	**Aggregate**	Fixed part	Variable part	Aggregate	Fixed part	Variable part	**Aggregate**
Group 1	0.611	2.570	**1.072**	***	***	***	0.611	2.570	**1.072**
Group 2	0.612	2.395	**0.906**	0.643	5.279	**1.001**	0.628	3.283	**0.952**
All	0.612	2.497	**0.988**	0.643	5.279	**1.001**	0.623	2.934	**0.992**

Contribution matrix									
	EP			A			EP / EPA (consolidated)		
	Fixed part	Variable part	**Aggregate**	Fixed part	Variable part	**Aggregate**	Fixed part	Variable part	**Aggregate**
Group 1	0.156	0.201	**0.357**				0.156	0.201	**0.357**
Group 2	0.175	0.135	**0.310**	0.193	0.132	**0.325**	0.367	0.268	**0.635**
All	0.331	0.336	**0.667**	0.193	0.132	**0.325**	0.523	0.469	**0.992**

$$\gamma_2 = \frac{R_2}{C_2} = \frac{R_0^2 + R_v^2}{C_0^2 + C_v^2} = \frac{C_0^2}{C_0^2 + C_v^2} \times \frac{R_0^2}{C_0^2} + \frac{C_v^2}{C_0^2 + C_v^2} \times \frac{R_v^2}{C_v^2}$$
$$= \frac{n_2\left(\frac{CF_{EP}}{n} + \frac{CF_A}{n_2}\right) + cQ_0^2}{n_2\left(\frac{CF_{EP}}{n} + \frac{CF_A}{n_2}\right) + cQ_2} \times \gamma_0^2 + \frac{c\left(Q_2 - Q_0^2\right)}{n_2\left(\frac{CF_{EP}}{n} + \frac{CF_A}{n_2}\right) + cQ_2} \times \gamma_v^2$$

the breakdown of coverage rates calculated for subscriber groups G1 and G2. By replacing γ_1 and γ_2 by their expressions in (11.30), we get:

$$\gamma = \frac{C_1^0}{C_1 + C_2} \times \gamma_0^1 + \frac{C_1^v}{C_1 + C_2} \times \gamma_v^1 + \frac{C_0^2}{C_1 + C_2} \times \gamma_0^2 + \frac{C_v^2}{C_1 + C_2} \times \gamma_v^2$$
$$= \frac{C_1^0}{C} \times \gamma_0^1 + \frac{C_0^2}{C} \times \gamma_0^2 + \frac{C_1^v}{C} \times \gamma_v^1 + \frac{C_v^2}{C} \times \gamma_v^2 \qquad (11.31)$$

This relationship states that the coverage rate of the service, $\gamma = R / C$, is a weighted average of the coverage rates of the captive and variable revenues of groups G1 and G2. This identifies 4 contributions (in percentage points) which are shown in a specific table (see the 4 cells coloured red in Table 52, page 184).

Numerical example Using the data from Table 52, we have :

$$0.998 = \frac{25.5}{100} \times 0.611 + \frac{58.5}{100} \times 0.628 + \frac{7.8}{100} \times 2.570 + \frac{8.1}{100} \times 3.283 = 0.156 + 0.367 + 0.201 + 0.268$$

Literally:

- the captive consumption of group G1 contributes 15.6% of the cost of the (general) service, with a self-financing ratio of 61.1%;
- captive consumption by group G2 contributes 36.7% of the cost of the service (general), with a self-financing ratio of 62.8%;
- the economic component of consumption in group G1 contributes 20.1% of the cost of the (general) service, with a multiplier of 2,570 ;
- the economic component of consumption in group G2 contributes 26.8% of the cost of the (general) service, with a multiplier of 3.283;

and there is a (small) deficit of 0.8% (as a proportion of the cost of the service) which is financed from other sources (necessarily). In addition,

- contributions from captive consumption are lower by $0.255 - 0.156 = 9.9$ percentage points and $0.585 - 0.367 = 21.8$ percentage points respectively,
- contributions on the variable part of consumption are higher by $0.201 - 0.078 = 12.3$ percentage points and $0.268 - 0.081 = 18.7$ percentage points

compared with their IBT-TBSE values. The relatively higher values for group 2 are due to the fact that the consumption of group 2 members is subsidised and taxed on an EPA service that includes the supply of drinking water and the supply of wastewater (whereas the consumption of group G1 members is subsidised and taxed on the drinking water service alone).

2) Service EP vs. Service A decomposition For the Service EP vs. Service A decomposition, we have:

$$\gamma = \frac{R}{C} = \frac{R_{\mathrm{EP}} + R_{\mathrm{A}}}{C_{\mathrm{EP}} + C_{\mathrm{A}}} = \frac{C_{\mathrm{EP}}}{C_{\mathrm{EP}} + C_{\mathrm{A}}} \times \frac{R_{\mathrm{EP}}}{C_{\mathrm{EP}}} + \frac{C_{\mathrm{A}}}{C_{\mathrm{EP}} + C_{\mathrm{A}}} \times \frac{R_{\mathrm{A}}}{C_{\mathrm{A}}} = \frac{C_{\mathrm{EP}}}{C_{\mathrm{EP}} + C_{\mathrm{A}}} \times \gamma_{\mathrm{EP}} + \frac{C_{\mathrm{A}}}{C_{\mathrm{EP}} + C_{\mathrm{A}}} \times \gamma_{\mathrm{A}} \qquad (11.32)$$

with:

- R_{EP} the Operator revenue (excluding taxes and environmental charges) generated by the marketing of the drinking water service (for the household subscriber population as a whole),

- R_{A} the Operator revenue (excluding taxes and environmental charges) generated by the marketing of the wastewater service (for the G2 group only, i.e. the sub-population of household subscribers connected to the wastewater network),

- $C_{\mathrm{EP}} = CF_{\mathrm{EP}} + c_{\mathrm{EP}} \times Q$ the cost of providing the drinking water service,

- $C_{\mathrm{A}} = CF_{\mathrm{A}} + c_{\mathrm{A}} \times Q_{\mathrm{A}}$ the cost of providing the wastewater treatment service,

and:

$$\gamma_{\mathrm{EP}} = \frac{R_{\mathrm{EP}}}{C_{\mathrm{EP}}} = \frac{C_0^{\mathrm{EP}}}{C_{\mathrm{EP}}} \times \frac{R_0^{\mathrm{EP}}}{C_0^{\mathrm{EP}}} + \frac{C_v^{\mathrm{EP}}}{C_{\mathrm{EP}}} \times \frac{R_v^{\mathrm{EP}}}{C_v^{\mathrm{EP}}} = \frac{C_0^{\mathrm{EP}}}{C_{\mathrm{EP}}} \times \gamma_0^{\mathrm{EP}} + \frac{C_v^{\mathrm{EP}}}{C_{\mathrm{EP}}} \times \gamma_v^{\mathrm{EP}}$$

$$\gamma_{\mathrm{A}} = \frac{R_{\mathrm{A}}}{C_{\mathrm{A}}} = \frac{C_0^{\mathrm{A}}}{C_{\mathrm{A}}} \times \frac{R_0^{\mathrm{A}}}{C_0^{\mathrm{A}}} + \frac{C_v^{\mathrm{A}}}{C_{\mathrm{A}}} \times \frac{R_v^{\mathrm{A}}}{C_v^{\mathrm{A}}} = \frac{C_0^{\mathrm{A}}}{C_{\mathrm{A}}} \times \gamma_0^{\mathrm{A}} + \frac{C_v^{\mathrm{A}}}{C_{\mathrm{A}}} \times \gamma_v^{\mathrm{A}}$$

the breakdown of the coverage rate for the drinking water and wastewater services respectively. By replacing γ_{EP} et γ_{A} by their expressions in (11.30), we get:

$$\begin{aligned}\gamma &= \frac{C_0^{\mathrm{EP}}}{C_1 + C_2} \times \gamma_0^{\mathrm{EP}} + \frac{C_v^{\mathrm{EP}}}{C_1 + C_2} \times \gamma_v^{\mathrm{EP}} + \frac{C_0^{\mathrm{A}}}{C_1 + C_2} \times \gamma_0^{\mathrm{A}} + \frac{C_v^{\mathrm{A}}}{C_1 + C_2} \times \gamma_v^{\mathrm{A}} \\ &= \frac{C_0^{\mathrm{EP}}}{C} \times \gamma_0^{\mathrm{EP}} + \frac{C_v^{\mathrm{EP}}}{C} \times \gamma_v^{\mathrm{EP}} + \frac{C_0^{\mathrm{A}}}{C} \times \gamma_0^{\mathrm{A}} + \frac{C_v^{\mathrm{A}}}{C} \times \gamma_v^{\mathrm{A}}\end{aligned} \qquad (11.33)$$

This relationship states that the coverage rate of the cost of general service, $\gamma = R / C$, is a weighted average of the coverage rates linked to the fixed part, in the economic sense, and to the variable part, in the economic sense, of the costs of the drinking water service and the costs of the wastewater service. This identifies 4 contributions (in percentage points) which are shown in a specific table (see the 4 cells coloured green in Table 52, page 184).

Numerical example Using the data from Table 52, we have :

$$0.998=\frac{54.1}{100}\times 0.612+\frac{13.5}{100}\times 2.497+\frac{30.0}{100}\times 0.643+\frac{2.5}{100}\times 5.279=0.331+0.336+0.193+0.132$$

Literally:

- captive consumption for the EP service contributes 33.1% of the cost of the general service EP/EPA, with a self-financing ratio of 61.2%,
- economic consumption for the EP service contributes 33.6% of the cost of the general service EP/EPA, with a multiplier of 2.497,
- captive consumption for service A contributes 19.3% of the cost of the general service EP/EPA, with a self-financing ratio of 64.3%,
- economic consumption for service A contributes 13.2% of the cost of the general service EP/EPA, with a multiplier of 5.279,

and there is a (small) deficit of 0.8% (as a proportion of the cost of the general service) which is financed from other sources (necessarily). In addition,

- contributions linked to the marketing of captive consumption are lower by $0.541-0.331=21.0$ percentage points for service EP, by $0.300-0.193=10.7$ percentage points for service A,
- contributions linked to the marketing of the variable part of consumption are higher by $0.336-0.135=20.1$ percentage points for service EP, by $0.132-0.025=10.7$ percentage points for service A,

compared with their IBT-TBSE values.

3) Decomposition EP-G1, EP-G2 and A-G2 Decompositions can be taken a step further by showing (i) for terms relating to group G2 in (11.31) decomposition Service EP vs. Service A with:

$$\frac{C_0^2}{C_1+C_2}\times\gamma_0^2=\frac{C_0^2}{C_1+C_2}\times\frac{R_0^2}{C_0^2}=\frac{C_0^2}{C_1+C_2}\times\frac{R_{0,2}^{\text{EP}}+R_{0,2}^{\text{A}}}{C_{0,2}^{\text{EP}}+C_{0,2}^{\text{A}}}=\frac{C_{0,2}^{\text{EP}}}{C_1+C_2}\times\frac{R_{0,2}^{\text{EP}}}{C_{0,2}^{\text{EP}}}+\frac{C_{0,2}^{\text{A}}}{C_1+C_2}\times\frac{R_{0,2}^{\text{A}}}{C_{0,2}^{\text{A}}}$$
$$=\frac{C_{0,2}^{\text{EP}}}{C_1+C_2}\times\gamma_{0,2}^{\text{EP}}+\frac{C_{0,2}^{\text{A}}}{C_1+C_2}\times\gamma_{0,2}^{\text{A}}$$

$$\frac{C_v^2}{C_1+C_2}\times\gamma_v^2=\frac{C_v^2}{C_1+C_2}\times\frac{R_v^2}{C_0^2}=\frac{C_v^2}{C_1+C_2}\times\frac{R_{v,2}^{\text{EP}}+R_{v,2}^{\text{A}}}{C_{v,2}^{\text{EP}}+C_{v,2}^{\text{A}}}=\frac{C_{v,2}^{\text{EP}}}{C_1+C_2}\times\frac{R_{v,2}^{\text{EP}}}{C_{v,2}^{\text{EP}}}+\frac{C_{v,2}^{\text{A}}}{C_1+C_2}\times\frac{R_{v,2}^{\text{A}}}{C_{v,2}^{\text{A}}}$$
$$=\frac{C_{v,2}^{\text{EP}}}{C_1+C_2}\times\gamma_{v,2}^{\text{EP}}+\frac{C_{v,2}^{\text{A}}}{C_1+C_2}\times\gamma_{v,2}^{\text{A}}$$

to obtain:

$$\gamma=\frac{C_1^0}{C_1+C_2}\times\gamma_0^1+\frac{C_1^v}{C_1+C_2}\times\gamma_v^1+\frac{C_{0,2}^{\text{EP}}}{C_1+C_2}\times\gamma_{0,2}^{\text{EP}}+\frac{C_{0,2}^{\text{A}}}{C_1+C_2}\times\gamma_{0,2}^{\text{A}}+\frac{C_{v,2}^{\text{EP}}}{C_1+C_2}\times\gamma_{v,2}^{\text{EP}}+\frac{C_{v,2}^{\text{A}}}{C_1+C_2}\times\gamma_{v,2}^{\text{A}}$$

(11.34)

The point is that this decomposition also corresponds to the one obtained by showing (ii) the distinction between groups G1 vs. G2 in the terms relating to the EP service in the decomposition (11.32) by making:

$$\frac{C_0^{\mathrm{EP}}}{C}\times\frac{R_0^{\mathrm{EP}}}{C_{\mathrm{EP}}}=\frac{C_0^{\mathrm{EP}}}{C_{\mathrm{EP}}+C_{\mathrm{A}}}\times\frac{R_{0,1}^{\mathrm{EP}}+R_{0,2}^{\mathrm{EP}}}{C_{0,1}^{\mathrm{EP}}+C_{0,2}^{\mathrm{EP}}}=\frac{C_{0,1}^{\mathrm{EP}}}{C_{\mathrm{EP}}+C_{\mathrm{A}}}\times\frac{R_{0,1}^{\mathrm{EP}}}{C_{0,1}^{\mathrm{EP}}}+\frac{C_{0,2}^{\mathrm{EP}}}{C_{\mathrm{EP}}+C_{\mathrm{A}}}\times\frac{R_{0,2}^{\mathrm{EP}}}{C_{0,2}^{\mathrm{EP}}}$$
$$=\frac{C_{0,1}^{\mathrm{EP}}}{C_{\mathrm{EP}}+C_{\mathrm{A}}}\times\gamma_{0,1}^{\mathrm{EP}}+\frac{C_{0,2}^{\mathrm{EP}}}{C_{\mathrm{EP}}+C_{\mathrm{A}}}\times\gamma_{0,2}^{\mathrm{EP}}$$

$$\frac{C_v^{\mathrm{EP}}}{C}\times\frac{R_v^{\mathrm{EP}}}{C_v^{\mathrm{EP}}}=\frac{C_v^{\mathrm{EP}}}{C_{\mathrm{EP}}+C_{\mathrm{A}}}\times\frac{R_{v,1}^{\mathrm{EP}}+R_{v,2}^{\mathrm{EP}}}{C_{v,1}^{\mathrm{EP}}+C_{v,2}^{\mathrm{EP}}}=\frac{C_{v,1}^{\mathrm{EP}}}{C_{\mathrm{EP}}+C_{\mathrm{A}}}\times\frac{R_{v,1}^{\mathrm{EP}}}{C_{v,1}^{\mathrm{EP}}}+\frac{C_{v,2}^{\mathrm{EP}}}{C_{\mathrm{EP}}+C_{\mathrm{A}}}\times\frac{R_{v,2}^{\mathrm{EP}}}{C_{v,2}^{\mathrm{EP}}}$$
$$=\frac{C_{v,1}^{\mathrm{EP}}}{C_{\mathrm{EP}}+C_{\mathrm{A}}}\times\gamma_{v,1}^{\mathrm{EP}}+\frac{C_{v,2}^{\mathrm{EP}}}{C_{\mathrm{EP}}+C_{\mathrm{A}}}\times\gamma_{v,2}^{\mathrm{EP}}$$

to obtain:

$$\gamma=\frac{C_{0,1}^{\mathrm{EP}}}{C_{\mathrm{EP}}+C_{\mathrm{A}}}\times\gamma_{0,1}+\frac{C_{0,2}^{\mathrm{EP}}}{C_{\mathrm{EP}}+C_{\mathrm{A}}}\times\gamma_{0,2}^{\mathrm{EP}}+\frac{C_0^{\mathrm{A}}}{C_{\mathrm{EP}}+C_{\mathrm{A}}}\times\gamma_0^{\mathrm{A}}+\frac{C_{v,1}^{\mathrm{EP}}}{C_{\mathrm{EP}}+C_{\mathrm{A}}}\times\gamma_{v,1}^{\mathrm{EP}}+\frac{C_{v,2}^{\mathrm{EP}}}{C_{\mathrm{EP}}+C_{\mathrm{A}}}\times\gamma_{v,2}^{\mathrm{EP}}+\frac{C_v^{\mathrm{A}}}{C_{\mathrm{EP}}+C_{\mathrm{A}}}\times\gamma_v^{\mathrm{A}}$$

given that $\gamma_{0,1}=\gamma_{0,1}^{\mathrm{EP}}$, $\gamma_{v,1}=\gamma_{v,1}^{\mathrm{EP}}$, $\gamma_0^{\mathrm{A}}=\gamma_{0,2}^{\mathrm{A}}$, $\gamma_v^{\mathrm{A}}=\gamma_{v,2}^{\mathrm{A}}$ and $C_1+C_2=C_{\mathrm{EP}}+C_{\mathrm{A}}=C$. See the 4 cells coloured in yellow in Table 52, page 184.

Numerical example Returning to (11.34), we have:

$$0.992=\frac{25.5}{100}\times0.611+\frac{28.6}{100}\times0.612+\frac{30}{100}\times0.643+\frac{7.8}{100}\times2.570+\frac{5.6}{100}\times2.395+\frac{2.5}{100}\times5.279$$
$$=0.156+0.175+0.135+0.201+0.193+0.132$$

Literally:

- the marketing of captive consumption by Group 1 for the EP service contributes 15.6% of the cost of the general service EP/EPA, with a self-financing ratio of 61.1%,
- the marketing of captive consumption by Group 2 contributes 17.5% of the cost of the general service for the EP component, with a self-financing ratio of 61.2%, and 19.3% for the A component, with a self-financing ratio of 64.4%,
- the marketing of the variable part of Group 1 consumption (for the EP service) contributes 20.1% of the cost of the general service EP/EPA, with a multiplier of 2.570,
- the marketing of the variable part of Group 2 consumption contributes 13.5% of the cost of the general service for the EP component, with a multiplier of 2.395, and 13.2% for the A component, with a multiplier of 5.279,

and there is a (small) deficit of 0.8% (as a proportion of the cost of the general service EP/EPA) which is financed from other sources (necessarily). See Table 52, page 184

B) Disaggregation of direct support Returning to the equation that breakdowns the coverage rate of the general service:

$$\frac{R}{C}=\frac{R_0+R_v}{C_0+C_v}=\frac{C_0}{C_0+C_v}\times\frac{R_0}{C_0}+\frac{C_v}{C_0+C_v}\times\frac{R_v}{C_0+C_v}=\frac{CF+cQ_0}{C}\times\gamma_0+\frac{c\left(Q-Q_0\right)}{C}\times\gamma_v$$

direct support is shown by:

$$\begin{aligned}1&=\frac{CF+cQ_0}{C}\times\gamma_0+\frac{c\left(Q-Q_0\right)}{C}\times\gamma_v+\frac{-\left(R-C\right)}{C}\\&=\frac{CF+cQ_0}{C}\times\gamma_0+\frac{c\left(Q-Q_0\right)}{C}\times\gamma_v+\frac{-\Pi}{C}\end{aligned}$$

The $-\Pi/C$ term breaks down as follows:

$$\frac{-\Pi}{C}=\frac{-\Pi_1-\Pi_2}{C_1+C_2}=\frac{C_1}{C_1+C_2}\times\frac{-\Pi_1}{C_1}+\frac{C_2}{C_1+C_2}\times\frac{-\Pi_2}{C_2}=\frac{C_1}{C}\times\mathrm{SD}_1+\frac{C_2}{C}\times\mathrm{SD}_2$$

$$\frac{-\Pi}{C}=\frac{-\Pi_{\mathrm{EP}}-\Pi_{\mathrm{A}}}{C_{\mathrm{EP}}+C_{\mathrm{A}}}=\frac{C_{\mathrm{EP}}}{C_{\mathrm{EP}}+C_{\mathrm{A}}}\times\frac{-\Pi_{\mathrm{EP}}}{C_{\mathrm{EP}}}+\frac{C_{\mathrm{A}}}{C_{\mathrm{EP}}+C_{\mathrm{A}}}\times\frac{-\Pi_{\mathrm{A}}}{C_{\mathrm{A}}}=\frac{C_{\mathrm{EP}}}{C}\times\mathrm{SD}_{\mathrm{EP}}+\frac{C_{\mathrm{A}}}{C}\times\mathrm{SD}_{\mathrm{A}}$$

$$\begin{aligned}\frac{-\Pi}{C}&=\frac{-\Pi_1^{\mathrm{EP}}-\Pi_2^{\mathrm{EP}}-\Pi_2^{\mathrm{A}}}{C_1^{\mathrm{EP}}+C_2^{\mathrm{EP}}+C_2^{\mathrm{A}}}=\frac{C_1^{\mathrm{EP}}}{C}\times\frac{-\Pi_1^{\mathrm{EP}}}{C_1^{\mathrm{EP}}}+\frac{C_2^{\mathrm{EP}}}{C}\times\frac{-\Pi_2^{\mathrm{EP}}}{C_1^{\mathrm{EP}}}+\frac{C_2^{\mathrm{A}}}{C}\times\frac{-\Pi_2^{\mathrm{A}}}{C_2^{\mathrm{A}}}\\&=\frac{C_1^{\mathrm{EP}}}{C}\times\mathrm{SD}_{\mathrm{EP\text{-}1}}+\frac{C_2^{\mathrm{EP}}}{C}\times\mathrm{SD}_{\mathrm{EP\text{-}2}}+\frac{C_2^{\mathrm{A}}}{C}\times\mathrm{SD}_{\mathrm{A\text{-}2}}\end{aligned}$$

(the notation "SD" for "Support Direct" in French) depending on whether disaggregating by groups, by services or by groups and services is performed.

Numerical example See Table 53, page 190, where:

$$\frac{-\Pi}{C}=\frac{33.3}{100}\times\left(-0.072\right)+\frac{66.7}{100}\times0.048=-0.024+0.032=0.08$$

in the case of group decomposition (with $\mathrm{SD}_1=-2.4\%$ et $\mathrm{SD}_2=+3.2\%$; see the 2 cells coloured red) and:

$$\frac{-\Pi}{C}=\frac{67.5}{100}\times0.012+\frac{32.5}{100}\times\left(-0.001\right)=0.008-0.0003=0.08$$

in the case of the breakdown by services (avec $\mathrm{SD}_{\mathrm{EP}}=+0.8\%$ et $\mathrm{SD}_{\mathrm{A}}=-0.03\%$; see the 2 cells coloured green) and :

$$\frac{-\Pi}{C}=\frac{33.3}{100}\times\left(-0.072\right)+\frac{34.2}{100}\times0.094+\frac{32.5}{100}\times\left(-0.001\right)=-0.024+0.032-0.0003=0.08$$

in the case of the breakdown by groups and services (with $\mathrm{SD}_{\mathrm{EP\text{-}1}}=-2.4\%$, $\mathrm{SD}_{\mathrm{EP\text{-}2}}=+3.2\%$ and $\mathrm{SD}_{\mathrm{A\text{-}2}}=-0.03\%$; see the 3 cells coloured yellow). Literally:

Table 53 : Breakdown of direct support by groups, factors, and groups and factors -- Summary tables

Service Cost									
	…	EP	Direct Support	…	A	Direct Support	…	EP EPA	Direct Support
Group 1	…	1 874 686.48 €	-134 647.68 €	…			…	1 874 686.48 €	-134 647.68 €
Group 2	…	1 924 915.97 €	180 917.47 €	…	1 826 749.43 €	-1 407.56 €	…	3 751 665.40 €	179 509.91 €
All	…	3 799 602.45 €	46 269.79 €	…	1 826 749.43 €	-1 407.56 €	…	5 626 351.88 €	44 862.23 €

Cost share matrix (Weights)									
	…		Total EP		…	Total A		…	Total EP / EPA
Group 1	…		0.333		…	***		…	0.333
Group 2	…		0.342		…	0.325		…	0.667
All	…		0.675		…	0.325		…	1.000

Coverage rate matrix									
	…	EP	Direct Support	…	A	Direct Support	…	EP EPA	Direct Support
Group 1	…	1.072	-0.072	…	***	***	…	1.072	-0.072
Group 2	…	0.906	0.094	…	1.001	-0.001	…	0.952	0.048
All	…	0.988	0.012	…	1.001	-0.001	…	0.992	0.008

Contribution matrix									
	…	EP	Direct Support	…	A	Direct Support	…	EP EPA	Direct Support
Group 1	…	0.357	-0.024	…	***	***	…	0.357	-0.024
Group 2	…	0.310	0.032	…	0.325	-0.00025	…	0.635	0.032
All	…	0.667	0.008	…	0.325	-0.00025	…	0.992	0.008

(1) Breakdown by group - Taking into account other categories of expenditure and revenue:

- households in group 1 are marked up on their EP consumption and households in group 2 are subsidised on their EPA consumption;
- The margins achieved on households in group G1, linked to the marketing of the EP service, contribute 2.4% to the financing of the general service with a mark-up of 7.2%;
- subsidies paid to households in group G2, linked to the marketing of the EPA service, represent 3.2% of the cost of the general service with a support rate of 4.8%;
- the cross-subsidy system is (in this case) slightly unbalanced, with a delta of $-2.4+3.2=0.8\%$ in the cost of the general service,

which feeds a deficit of the same amount (in percentage points) for the financing of the general service, which is then financed by other sources (necessarily).

(2) Breakdown by services - Taking into account other categories of expenditure and revenue:

- direct support for the EP service is performed, with a support rate of 1.2% (of the cost of the EP service, in relation to the other categories of expenditure and revenue), which accounts for 0.8% of the cost of the (general) service;
- direct taxation is levied on service A, with a contribution rate of 0.1% (of the cost of service A, linked to the other categories of expenditure and revenue), which represents 0.03% of the cost of the (general) service;

and the mass of subsidies (on EP service) being here greater than the mass of contributions (on the A service):

- a slight deficit for the general service, $0.008+(-0.0003)=0.77\%$ of the cost of the general service, is recorded with a loss-making on EP service (0.8% of the total cost) and an A service that is (almost) in balance (the margin generated on the A represents 0.03% of the total cost).

This (small) deficit is then (necessarily) financed by other sources.

(3) Breakdown by group and service - Taking into account other categories of expenditure and revenue:

- Households in Group 1 are ultimately margined on their consumption, with a mark-up of 7.2%, and contribute 2.4% to the financing of the general service via these margins;
- Households in group G2 are ultimately (i) subsidised on the provision of the EP service with a support rate of 9.4% for a (total) amount representing 3.2% of the cost of the general service, (ii) margined on the provision of service A, with a mark-up of 1 per 1000 and contribute via these margins to the financing of the general service up to 0.025%, (iii) subsidised on the EPA service with a support rate of 4.8% for a (total) amount representing 3.2% of the cost of the (general) service.

As the margins generated on households in group G1, linked to the marketing of the EP service alone, are lower than the subsidies paid in fine to households in group G2, via the marketing of the EPA service, the result is a financing requirement, amounting to $-2.4+3.2-0.025=0.8\%$ of the cost of the general service, which is then covered by other sources of financing (necessarily).

C) Additional disaggregation of relationships (11.27) & (11.29) Using similar calculations, one gets (without additional difficulty) the disaggregation by groups (G1 vs. G2), by services (EP vs. A), and by groups and services (G1-EP, G2-EP, G2) of relationships (11.27) and (11.29). We get:

(1) As regards the term $\frac{\underline{C}}{C}\times\underline{\gamma}$ relating to (probable) basic support :

a) Disaggregation by groups (G1 vs. G2) :

$$\frac{C(\underline{Q})}{C(Q)}\times\underline{\gamma}=\frac{\underline{C}_1}{C}\times\underline{\gamma}_1+\frac{\underline{C}_2}{C}\times\underline{\gamma}_2$$

with $\underline{\gamma}_1=\frac{\underline{R}_1}{\underline{C}_1}$ and $\underline{\gamma}_2=\frac{\underline{R}_2}{\underline{C}_2}$;

b) Disaggregation by Services (EP vs. A):

$$\frac{C(\underline{Q})}{C(Q)}\times\underline{\gamma}=\frac{\underline{C}_{\mathrm{EP}}}{C}\times\underline{\gamma}_{\mathrm{EP}}+\frac{\underline{C}_{\mathrm{A}}}{C}\times\underline{\gamma}_{\mathrm{A}}$$

with $\underline{\gamma}_{\mathrm{EP}}=\frac{\underline{R}_{\mathrm{EP}}}{\underline{C}_{\mathrm{EP}}}$ and $\underline{\gamma}_{\mathrm{A}}=\frac{\underline{R}_{\mathrm{A}}}{\underline{C}_{\mathrm{A}}}$;

c) Disaggregation by Groups and Services (EP-G1, EP-G2 and A-G2) :

$$\frac{C(\underline{Q})}{C(Q)}\times\underline{\gamma}=\frac{\underline{C}_1^{\mathrm{EP}}}{C}\times\underline{\gamma}_1^{\mathrm{EP}}+\frac{\underline{C}_2^{\mathrm{EP}}}{C}\times\underline{\gamma}_2^{\mathrm{EP}}+\frac{\underline{C}_2^{\mathrm{A}}}{C}\times\underline{\gamma}_2^{\mathrm{A}}$$

with $\underline{\gamma}_1^{\mathrm{EP}}=\frac{\underline{R}_1^{\mathrm{EP}}}{\underline{C}_1^{\mathrm{EP}}}$, $\underline{\gamma}_2^{\mathrm{EP}}=\frac{\underline{R}_2^{\mathrm{EP}}}{\underline{C}_2^{\mathrm{EP}}}$ and $\underline{\gamma}_2^{\mathrm{A}}=\frac{\underline{R}_2^{\mathrm{A}}}{\underline{C}_2^{\mathrm{A}}}$.

(2) As regards the term relating to captive but non-basic consumption:

a) Disaggregation by groups (G1 vs. G2) :

$$\frac{c(Q_0-\underline{Q})}{C(Q)}\times\gamma_{\mathrm{Fixe\_HB}}=\frac{C_{0,1}-\underline{C}_1}{C(Q)}\times\gamma_{\mathrm{Fixe\_HB,1}}+\frac{C_{0,2}-\underline{C}_2}{C(Q)}\times\gamma_{\mathrm{Fixe\_HB,2}}$$

with $\gamma_{\mathrm{Fixe\_HB,1}}=\frac{R_{0,1}-\underline{R}_1}{C_{0,1}-\underline{C}_1}$ and $\gamma_{\mathrm{Fixe\_HB,2}}=\frac{R_{0,2}-\underline{R}_2}{C_{0,2}-\underline{C}_2}$;

b) Disaggregation by Services (EP vs. A):

$$\frac{c_{\mathrm{EP}}(Q_0-\underline{Q})+c_{\mathrm{A}}(Q_0^{\mathrm{A}}-\underline{Q}_{\mathrm{A}})}{C(Q)}\times\gamma_{\mathrm{Fixe\_HB}}=\frac{C_0^{\mathrm{EP}}-\underline{C}_{\mathrm{EP}}}{C(Q)}\times\gamma_{\mathrm{Fixe\_HB}}^{\mathrm{EP}}+\frac{C_0^{\mathrm{A}}-\underline{C}_{\mathrm{A}}}{C(Q)}\times\gamma_{\mathrm{Fixe\_HB}}^{\mathrm{A}}$$

with $\gamma_{\text{Fixe\_HB}}^{\text{EP}} = \dfrac{R_0^{\text{EP}} - \underline{R}_{\text{EP}}}{C_0^{\text{EP}} - \underline{C}_{\text{EP}}}$ and $\gamma_{\text{Fixe\_HB}}^{\text{A}} = \dfrac{R_0^{\text{A}} - \underline{R}_{\text{EP}}}{C_0^{\text{A}} - \underline{C}_{\text{A}}}$;

c) Disaggregation by Groups and Services (EP-G1, EP-G2 and A-G2) :

$$\frac{c_{\text{EP}}\left(Q_0 - \underline{Q}\right) + c_{\text{A}}\left(Q_0^{\text{A}} - \underline{Q}_{\text{A}}\right)}{C(Q)} \times \gamma_{\text{Fixe\_HB}} = \frac{C_{0,1}^{\text{EP}} - \underline{C}_{\text{EP}}^1}{C(Q)} \times \gamma_{\text{Fixe\_HB},1}^{\text{EP}} + \frac{C_{0,2}^{\text{EP}} - \underline{C}_{\text{EP}}^2}{C(Q)} \times \gamma_{\text{Fixe\_HB},2}^{\text{EP}} + \frac{C_{0,2}^{\text{A}} - \underline{C}_{\text{A}}^2}{C(Q)} \times \gamma_{\text{Fixe\_HB},2}^{\text{A}}$$

with $\gamma_{\text{Fixe\_HB},1}^{\text{EP}} = \dfrac{R_{0,1}^{\text{EP}} - \underline{R}_1^{\text{EP}}}{C_{0,1}^{\text{EP}} - \underline{C}_{\text{EP}}^1}$, $\gamma_{\text{Fixe\_HB},2}^{\text{EP}} = \dfrac{R_{0,2}^{\text{EP}} - \underline{R}_2^{\text{EP}}}{C_{0,2}^{\text{EP}} - \underline{C}_{\text{EP}}^2}$ and $\gamma_{\text{Fixe\_HB},2}^{\text{A}} = \dfrac{R_{0,2}^{\text{A}} - \underline{R}_2^{\text{A}}}{C_{0,2}^{\text{A}} - \underline{C}_{\text{A}}^2}$.

(3) As regards the term relating to the economic part of consumption excluding overconsumptions:

a) Disaggregation by groups (G1 vs. G2) :

$$\frac{c_{\text{EP}}\left(Q_{\kappa=1} - Q_0\right) + c_{\text{A}}\left(Q_{\kappa=1}^{\text{A}} - Q_0^{\text{A}}\right)}{C(Q)} \times \gamma_v^{\kappa=1} = \frac{c_{\text{EP}}\left(Q_{\kappa=1,1} - Q_{0,1}\right)}{C(Q)} \times \gamma_{v,1}^{\kappa=1} + \frac{c_{\text{EPA}}\left(Q_{\kappa=1,2} - Q_{0,2}\right)}{C(Q)} \times \gamma_{v,2}^{\kappa=1}$$

with $c_{\text{EP}} + c_{\text{A}} = c_{\text{EPA}}$, $\gamma_{v,1}^{\kappa=1} = \dfrac{R_{v,1}^{\kappa=1} - R_{0,1}}{c_{\text{EP}}\left(Q_{\kappa=1,1} - Q_{0,1}\right)}$ and $\gamma_{v,2}^{\kappa=1} = \dfrac{R_{v,2}^{\kappa=1} - R_{0,2}}{c_{\text{EPA}}\left(Q_{\kappa=1,2} - Q_{0,2}\right)}$;

b) Disaggregation by Services (EP vs. A):

$$\frac{c_{\text{EP}}\left(Q_{\kappa=1} - Q_0\right) + c_{\text{A}}\left(Q_{\kappa=1}^{\text{A}} - Q_0^{\text{A}}\right)}{C(Q)} \times \gamma_v^{\kappa=1} = \frac{c_{\text{EP}}\left(Q_{\kappa=1} - Q_0\right)}{C(Q)} \times \gamma_{v,\text{EP}}^{\kappa=1} + \frac{c_{\text{A}}\left(Q_{\kappa=1}^{\text{A}} - Q_0^{\text{A}}\right)}{C(Q)} \times \gamma_{v,\text{A}}^{\kappa=1}$$

with $Q_{\kappa=1,2} = Q_{\kappa=1}^{\text{A}}$, $\gamma_{v,\text{EP}}^{\kappa=1} = \dfrac{R_{v,\text{EP}}^{\kappa=1} - R_{0,\text{EP}}}{c_{\text{EP}}\left(Q_{\kappa=1} - Q_0\right)}$ and $\gamma_{v,\text{A}}^{\kappa=1} = \dfrac{R_{v,\text{A}}^{\kappa=1} - R_{0,\text{A}}}{c\left(Q_{\kappa=1,2} - Q_{0,2}\right)}$;

c) Disaggregation by Groups and Services (EP-G1, EP-G2 and A-G2) :

$$\begin{aligned}\frac{c_{\text{EP}}\left(Q_{\kappa=1} - Q_0\right) + c_{\text{A}}\left(Q_{\kappa=1}^{\text{A}} - Q_0^{\text{A}}\right)}{C(Q)} \times \gamma_v^{\kappa=1} &= \frac{c_{\text{EP}}\left(Q_{\kappa=1,1} - Q_{0,1}\right)}{C(Q)} \times \gamma_{v,\text{EP},1}^{\kappa=1} \\ &+ \frac{c_{\text{EP}}\left(Q_{\kappa=1,2} - Q_{0,2}\right)}{C(Q)} \times \gamma_{v,\text{EP},1}^{\kappa=1} \\ &+ \frac{c_{\text{A}}\left(Q_{\kappa=1,2}^{\text{A}} - Q_{0,2}^{\text{A}}\right)}{C(Q)} \times \gamma_{v,\text{A},2}^{\kappa=1}\end{aligned}$$

with $\gamma_{v,\text{EP},1}^{\kappa=1} = \dfrac{R_{v,1}^{\kappa=1,\text{EP}} - R_{0,\text{EP}}^1}{c_{\text{EP}}\left(Q_{\kappa=1,1} - Q_{0,1}\right)}$, $\gamma_{v,\text{EP},2}^{\kappa=1} = \dfrac{R_{v,2}^{\kappa=1,\text{EP}} - R_{0,\text{EP}}^2}{c_{\text{EP}}\left(Q_{\kappa=1,2} - Q_{0,2}\right)}$ and $\gamma_{v,\text{A},2}^{\kappa=1} = \dfrac{R_{v,2}^{\kappa=1,\text{A}} - R_{0,\text{A}}^2}{c_{\text{A}}\left(Q_{\kappa=1,2} - Q_{0,2}\right)}$.

(4) As regards the term relating to over-consumption (linked to tariff misperception) :

a) Disaggregation by groups (G1 vs. G2) :

$$\frac{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0}-Q_{\kappa=1}\right)+c_{\mathrm{A}}\left(Q^{\mathrm{A}}_{\kappa=\kappa_0}-Q^{\mathrm{A}}_{\kappa=1}\right)}{C(Q)}\times\gamma_{\mathrm{surco}}=\frac{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0,1}-Q_{\kappa=1,1}\right)}{C(Q)}\times\gamma_{\mathrm{surco},1}+\frac{c_{\mathrm{EPA}}\left(Q_{\kappa=\kappa_0,2}-Q_{\kappa=1,2}\right)}{C(Q)}\times\gamma_{\mathrm{surco},2}$$

with $\gamma_{\mathrm{surco},1}=\dfrac{R^{\kappa=\kappa_0}_{v,1}-R^{\kappa=1}_{v,1}}{c_{\mathrm{EP}}\left(Q_{\kappa=1,1}-Q_{0,1}\right)}$ and $\gamma_{\mathrm{surco},2}=\dfrac{R^{\kappa=\kappa_0}_{v,2}-R^{\kappa=1}_{v,2}}{c_{\mathrm{EPA}}\left(Q_{\kappa=1,2}-Q_{0,2}\right)}$;

b) Disaggregation by Services (EP vs. A):

$$\frac{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0}-Q_{\kappa=1}\right)+c_{\mathrm{A}}\left(Q^{\mathrm{A}}_{\kappa=\kappa_0,1}-Q^{\mathrm{A}}_{\kappa=1}\right)}{C(Q)}\times\gamma_{\mathrm{surco}}=\frac{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0}-Q_{\kappa=1}\right)}{C(Q)}\times\gamma^{\mathrm{EP}}_{\mathrm{surco}}+\frac{c_{\mathrm{A}}\left(Q^{\mathrm{A}}_{\kappa=\kappa_0}-Q^{\mathrm{A}}_{\kappa=1}\right)}{C(Q)}\times\gamma^{\mathrm{A}}_{\mathrm{surco}}$$

with: $\gamma^{\mathrm{EP}}_{\mathrm{surco}}=\dfrac{R^{\kappa=\kappa_0}_{v,\mathrm{EP}}-R^{\kappa=1}_{v,\mathrm{EP}}}{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0}-Q_{\kappa=1}\right)}$ and $\gamma^{\mathrm{A}}_{\mathrm{surco}}=\dfrac{R^{\kappa=\kappa_0}_{v,\mathrm{A}}-R^{\kappa=1}_{v,\mathrm{A}}}{c_{\mathrm{A}}\left(Q^{\mathrm{A}}_{\kappa=\kappa_0}-Q^{\mathrm{A}}_{\kappa=1}\right)}$;

c) Disaggregation by Groups and Services (EP-G1, EP-G2 and A-G2) :

$$\begin{aligned}\frac{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0}-Q_{\kappa=1}\right)+c_{\mathrm{A}}\left(Q^{\mathrm{A}}_{\kappa=\kappa_0,1}-Q^{\mathrm{A}}_{\kappa=1}\right)}{C(Q)}\times\gamma_{\mathrm{surco}}&=\frac{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0,1}-Q_{\kappa=1,1}\right)}{C(Q)}\times\gamma^{\mathrm{EP}}_{\mathrm{surco},1}\\&+\frac{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0,2}-Q_{\kappa=1,2}\right)}{C(Q)}\times\gamma^{\mathrm{EP}}_{\mathrm{surco},2}\\&+\frac{c_{\mathrm{A}}\left(Q^{\mathrm{A}}_{\kappa=\kappa_0,2}-Q^{\mathrm{A}}_{\kappa=1,2}\right)}{C(Q)}\times\gamma^{\mathrm{A}}_{\mathrm{surco},2}\end{aligned}$$

with:

$\gamma^{\mathrm{EP}}_{\mathrm{surco},1}=\dfrac{R^{\kappa=\kappa_0,\mathrm{EP}}_{v,1}-R^{\kappa=1,\mathrm{EP}}_{v,1}}{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0,1}-Q_{\kappa=1,1}\right)}$ $\gamma^{\mathrm{EP}}_{\mathrm{surco},2}=\dfrac{R^{\kappa=\kappa_0,\mathrm{EP}}_{v,2}-R^{\kappa=1,\mathrm{EP}}_{v,2}}{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0,2}-Q_{\kappa=1,2}\right)}$ and $\gamma^{\mathrm{A}}_{\mathrm{surco},2}=\dfrac{R^{\kappa=\kappa_0,\mathrm{A}}_{v,2}-R^{\kappa=1,\mathrm{A}}_{v,2}}{c_{\mathrm{A}}\left(Q^{\mathrm{A}}_{\kappa=\kappa_0,2}-Q^{\mathrm{A}}_{\kappa=1,2}\right)}$.

These different contributions and their breakdowns are then displayed in specific tables, in addition to the breakdowns of direct support (by Groups (G1 vs. G2), by Services (EP vs. A) and by Groups and Services (EP-G1, EP-G2 and A-G2)) given above.

See the set of tables "Table 54" for an illustration.

Table 54 :

Service cost matrix

EP	Fixed part		Variable part			
	Basic	Captive non basic	PP Economic part	Overcomsumption	Total	Direct support
Group 1	1 364 962.36 €	69 131.58 €	381 585.69 €	59 006.85 €	1 874 686.48 €	-134 647.68 €
Group 2	1 546 422.65 €	61 050.69 €	276 631.61 €	40 811.03 €	1 924 915.97 €	180 917.47 €
All	2 911 385.02 €	130 182.26 €	658 217.29 €	99 817.88 €	3 799 602.45 €	46 269.79 €

A	Fixed part		Variable part			
	Basic	Captive non basic	PP Economic part	Overcomsumption	Total	Direct support
Group 1						
Group 2	1 658 530.17 €	27 133.64 €	122 947.38 €	18 138.23 €	1 826 749.43 €	-1 407.56 €
All	1 658 530.17 €	27 133.64 €	122 947.38 €	18 138.23 €	1 826 749.43 €	-1 407.56 €

Consolidated	Fixed part		Variable part			
	Basic	Captive non basic	PP Economic part	Overcomsumption	Total	Direct support
Group 1	1 364 962.36 €	69 131.58 €	381 585.69 €	59 006.85 €	1 874 686.48 €	-134 647.68 €
Group 2	3 204 952.83 €	88 184.33 €	399 578.99 €	58 949.26 €	3 751 665.40 €	179 509.91 €
All	4 569 915.19 €	157 315.90 €	781 164.67 €	117 956.11 €	5 626 351.88 €	44 862.23 €

Cost share matrix (weights)

EP	Fixed part		Variable part			
	Basic	Captive non basic	PP Economic part	Overcomsumption	Total	Direct support
Group 1	0.243	0.012	0.068	0.010	0.333	
Group 2	0.275	0.011	0.049	0.007	0.342	
All	0.517	0.023	0.117	0.018	0.675	

A	Fixed part		Variable part			
	Basic	Captive non basic	PP Economic part	Overcomsumption	Total	Direct support
Group 1						
Group 2	0.295	0.005	0.022	0.003	0.325	
All	0.295	0.005	0.022	0.003	0.325	

Consolidated	Fixed part		Variable part			
	Basic	Captive non basic	PP Economic part	Overcomsumption	Total	Direct support
Group 1	0.243	0.012	0.068	0.010	0.333	
Group 2	0.570	0.016	0.071	0.010	0.667	
All	0.812	0.028	0.139	0.021	1.000	

Coverage rate matrix

EP	Fixed part		Variable part			
	Basic	Captive non basic	PP Economic part	Overcomsumption	Total	Direct support
Group 1	0.552	1.777	2.470	3.218	1.072	-0.072
Group 2	0.562	1.880	2.326	2.863	0.906	0.094
All	0.557	1.825	2.410	3.073	0.988	0.012

A	Fixed part		Variable part			
	Basic	Captive non basic	PP Economic part	Overcomsumption	Total	Direct support
Group 1						
Group 2	0.572	4.934	5.247	5.495	1.001	-0.001
All	0.572	4.934	5.247	5.495	1.001	-0.001

Consolidated	Fixed part		Variable part			
	Basic	Captive non basic	PP Economic part	Overcomsumption	Total	Direct support
Group 1	0.552	1.777	2.470	3.218	1.072	-0.072
Group 2	0.567	2.819	3.225	3.673	0.952	0.048
All	0.563	2.361	2.856	3.446	0.992	0.008

Contribution matrix

EP	Fixed part		Variable part			
	Basic	Captive non basic	PP Economic part	Overcomsumption	Total	Direct support
Group 1	0.134	0.022	0.168	0.034	0.357	-0.024
Group 2	0.154	0.020	0.114	0.021	0.310	0.032
All	0.288	0.042	0.282	0.055	0.667	0.008

A	Fixed part		Variable part			
	Basic	Captive non basic	PP Economic part	Overcomsumption	Total	Direct support
Group 1						
Group 2	0.169	0.024	0.115	0.018	0.325	-0.0003
All	0.169	0.024	0.115	0.018	0.325	-0.0003

Consolidated	Fixed part		Variable part			
	Basic	Captive non basic	PP Economic part	Overcomsumption	Total	Direct support
Group 1	0.134	0.022	0.168	0.034	0.357	-0.024
Group 2	0.323	0.044	0.229	0.038	0.635	0.0319
All	0.457	0.066	0.397	0.072	0.992	0.0080

11.2.3 Basic service funding

The aim here is to identify (and quantify) the sources of funding for the Basic Service (which is presumably subsidised by the other components of demand). To this end, one proceeds as follows.

(1) One starts with the accounting identity :

$$1=\frac{CF+c\underline{Q}}{C(Q)}\times\underline{\gamma}+\frac{c\left(Q_0-\underline{Q}\right)}{C(Q)}\times\gamma_{\text{Fixe\_HB}}+\frac{c\left(Q_{\kappa=1}-Q_0\right)}{C(Q)}\times\gamma_{\text{v}}^{\kappa=1}+\frac{c\left(Q_{\kappa=\kappa_0}-Q_{\kappa=1}\right)}{C(Q)}\times\gamma_{\text{surco}}+\frac{-\Pi}{C(Q)}$$

showing (potential) direct support (in the event of an operating deficit $\Pi=R-C<0$), next

(2) We have:

$$-\frac{CF+c\underline{Q}}{C(Q)}\times\left(\underline{\gamma}-1\right)=\frac{c\left(Q_0-\underline{Q}\right)}{C(Q)}\times\left(\gamma_{\text{Fixe\_HB}}-1\right)+\frac{c\left(Q_{\kappa=1}-Q_0\right)}{C(Q)}\times\left(\gamma_{\text{v}}^{\kappa=1}-1\right)$$
$$+\frac{c\left(Q_{\kappa=\kappa_0}-Q_{\kappa=1}\right)}{C(Q)}\times\left(\gamma_{\text{surco}}-1\right)+\frac{-\Pi}{C(Q)}$$

and:

$$1=\underline{\gamma}+\frac{c\left(Q_0-\underline{Q}\right)}{\underline{C}}\times\left(\gamma_{\text{Fixe\_HB}}-1\right)+\frac{c\left(Q_{\kappa=1}-Q_0\right)}{\underline{C}}\times\left(\gamma_{\text{v}}^{\kappa=1}-1\right)+\frac{c\left(Q_{\kappa=\kappa_0}-Q_{\kappa=1}\right)}{\underline{C}}\times\left(\gamma_{\text{surco}}-1\right)+\frac{-\Pi}{\underline{C}}$$

(11.35)

with $\underline{C}=C\left(\underline{Q}\right)=CF+c\underline{Q}$ the cost to the general service EP/EPA of providing the basic service. This equation breaks down the funding of the basic service into its various components:

(1) self-financing,

(2) the captive but non-basic part of consumption (by default, the uses of water for garden and swimming pool maintenance),

(3) the so-called "economic" part of demand, excluding excess consumption,

(4) over-consumption (due to poor perception of the tariff)

and :

(5) (Potential) direct support (calculated by balance).

The contributions (expressed in percentage points; see below) of funding sources (2), (3) and (4) can then themselves be broken down ex post as the product of a volume effect (in fact, the equivalent of a volume) multiplied by an average margin rate, linked to the characteristics of demand, pricing policy and service cost.

Table 55 : Breakdown of basic service funding (in General Population)

	Sel-financing (%)	Support	Funding sources	Margin rate	Ratio u.s.e.	Contributions In percentage points %	Contributions in %
Basic service	56.3	**43.7**	Non basic captive part	1.361	0.0344	4.7	10.7
			PP Q ($\kappa = 1$)	1.856	0.1709	31.7	72.6
			Overconsumptions	2.446	0.0258	6.3	14.4
			Direct support			1.0	2.2
			Total			**43.7**	**100.0**

linked to Table 46 that writtes:

	Fixed Part				Variable part					
n	Mean	Mass	Mean	Mass	Mean	Mass	Mean	Mass	Mean	Mass
47847	Basic Consumption		Non basic Captive Consumption		PP Economic Part		Overconsumption			
Revenue	53.76	2572411.44	7.76	371433.87	46.63	2231219.78	8.49	406424.55	116.65	5581489.65
Cost	95.51	**4569915.19**	3.29	**157315.90**	16.33	**781164.67**	2.47	**117956.11**	117.59	5626351.88
Net margin	-41.75	-1997503.75	4.48	214117.97	30.31	1450055.11	6.03	288468.44	-0.94	-44862.23
Coverage rate		**0.563**		**2.361**		**2.856**		**3.446**		0.992
	Cost shares									
		Contribution								Sum
Breakdown	0.812	0.457	0.028	0.066	0.139	0.397	0.02	0.072		0.992

Numerical example See Table 55 (derived from the data of Table 46, page 176). The latter displays the following breakdown:

$$\begin{aligned} 1 &= 0.563 + \frac{0.028}{0.812} \times (2.361 - 1) + \frac{0.139}{0.812} \times (2.856 - 1) + \frac{0.02}{0.812} \times (3.446 - 1) + \frac{-(0.992 - 1)}{0.812} \\ &= 0.563 + \frac{1}{29} \times 1.361 + \frac{139}{812} \times 1.856 + \frac{1}{406} \times 2.446 + \frac{2}{203} \\ &= 0.563 + 0.047 + 0.318 + 0.006 + 0.009 \end{aligned}$$

Next:

- Coverage rates greater than unity indicate a source of funding for consumption components whose coverage rate is less than unity.

With regard to the values obtained, the data show :

- for every euro invested in the provision of basic service (consisting of the access to the network and basic consumption), 0.563 euros will be recovered in terms of direct revenue from the marketing of the basic service,

i.e. 56.3% of basic service provision is self-financed, with the balance $1 - 0.563 = 0.437 = 43.7\%$ that is financed by the marketing of other production components and, where applicable, direct support[48]. The following observations can then be made:

- production linked to captive but non-basic consumption finances the provision of the basic service up to 4.7 percentage points,
- excluding over-consumption, production linked to the variable part of consumption finances the provision of the basic service up to 31.7 percentage points,
- the contributions (to the funding of basic service) of overconsumption and direct support (operating deficit) are rather weak (less than 1 percentage point).

These contributions in percentage points can be reformulated as % of the loss on basic consumption (see the last column of Table 55) with the following observations:

- $\frac{4.7}{43.7} = 10.7\%$ of the Operator's deficit for the provision of the basic service is financed by the production linked to the satisfaction of captive but non basic consumption, that is here the water uses for garden maitenance and swimming pool maintenance,
- $\frac{31.7}{43.7} = 72.5\%$ of the Operator's deficit for the provision of the basic service is financed by the economic component of consumption excluding overconsumption,
- $\frac{0.6}{43.7} = 1.4\%$ of the Operator's deficit for the provision of the basic service is financed by overconsumption linked to tariff misperception;

and direct support (deficit) accounts for $\frac{0.9}{43.7} = 2.1\%$ of the Operator deficit linked to the provision of the basic service.

[48] This interpretation because the basic service is the only component of consumption for which the coverage rate is less than one.

INN WATER

Next, these contributions can be broken down as the product of a volume (in fact, the equivalent of a volume) multiplied by a margin rate. Thus, if focus is on the variable part of consumption excluding over-consumption, the contribution (to the financing of basic service) of the production linked to meet this type of consumption breaks down as follows:

$$\begin{aligned} 0.318 &= \frac{C(Q_{\kappa=1}) - C_0}{\underline{C}} \times \left(\gamma_v^{\kappa=1} - 1\right) \\ &= \frac{c \times (Q_{\kappa=1} - Q_0)}{CF + c \times \underline{Q}} \times \left(\frac{\sum T\left(q_i^{\kappa=1}\right) - T\left(q_{0,i}\right)}{c \times \left(Q_{\kappa=1} - Q_0\right)} - 1 \right) \\ &= \frac{139}{812} \times (2.856 - 1) = 0.171 \times 1.856 \end{aligned}$$

The secund factor:

$$\begin{aligned} \gamma_v^{\kappa=1} - 1 &= \frac{\sum T\left(q_i^{\kappa=1}\right) - T\left(q_{0,i}\right)}{c \times \left(Q_{\kappa=1} - Q_0\right)} - 1 \\ &= \frac{1}{c} \times \left(\frac{\sum T\left(q_i^{\kappa=1}\right) - T\left(q_{0,i}\right)}{Q_{\kappa=1} - Q_0} - c \right) \\ &= 2.856 - 1 = 1.856 \end{aligned}$$

is linked to the margin (average, per unit of service) generated by the production of a cubic metre intended to satisfy the economic needs, excluding over-consumption, of a household. More specifically:

- Excluding over-consumption, the production of a cubic metre to satisfy "economic" uses costs c euros and generates a margin of $2.856 \times c - c = 1.856 \times c$ euros (on average), which is then allocated to financing the provision of the basic service;

or, equivalently:

- Excluding over-consumption, the production of one cubic metre to satisfy "economic" uses produces (on average) a surplus of $2.856 - 1 = 1.856$ cubic metre which can be allocated to basic service supply.

At the same time,

- the characteristics of the tariff and the characteristics of demand generate economic consumption, excluding overconsumption, of $Q_{\kappa=1} - Q_0$ cubic metres,

and therefore:

- a production of $\left(Q_{\kappa=1} - Q_0\right) \times (2.856 - 1) = \left(Q_{\kappa=1} - Q_0\right) \times 1.856$ cubic metre that can be allocated to basic service supply.

Besides, the provision of basic service (which includes the access to the network) for a population of *n* household customers costs:

$C\left(\underline{Q}\right) = CF + c \times \underline{Q}$ euros

that is the equivalent of:

$$C\left(\underline{Q}\right) \times \frac{1}{c} = \left(CF + c\underline{Q}\right) \times \frac{1}{c} = \frac{CF}{c} + \underline{Q} \text{ service units (cubic metres)}$$

In this regard, the proportion of production, in cubic metre equivalent (service units), for the basic service supply which is covered by the part of production for the economic needs of households, excluding overconsumption, sets to :

$$\frac{Q_{\kappa=1} - Q_0}{\frac{CF}{c} + \underline{Q}} = \frac{0.318}{1.856} = 0.171 = 17.1\%$$

what corresponds to the value of the first factor. Similar reasoning leads to the following conclusions:

- non-basic captive consumption, that is water uses related to garden maintenance and swimming pool maintenance, which initially represented 3.5% of production intended for the supply of the basic service, in cubic metre equivalents, ultimately contributed 4.7% of production intended for basic service supply, in cubic metre equivalents, with a margin rate, linked to the marketing of these units, of $2.361 - 1 = 1.361$;

- overconsumption, linked to tariff misperception, which initially represented 2.5% of the production intended basic service supply, in cubic metre equivalents, ultimately contributed to 0.6% of the production intended for basic service supply, in cubic metre equivalents, with a margin rate, linked to the marketing of these units, of $3.446 - 1 = 2.446$.

Lastly, other sources related to direct support ultimately account for 0.9%, in cubic metre equivalents, of the production used to basic service supply.

Decomposition Similar to general service funding, the relationship which identifies and quantifies the sources of financing for basic service can be disaggregated by group, by service and by group and service. See:

- equation (11.36), on next page, and Table 56 (numerical example), page 205, for group desagregation (G1 vs. G2) ;
- equation (11.37), on next page, and Table 57 (numerical example), page 206, for Service desagregation (EP vs. A) ;
- equation (11.38), on next page, and Table 58 (numerical example), page 207, for Group and Service disaggregation (EP-G1, EP-and A-G2)

and Appendix 8 for the derivation of these accounting relationships.

.

Group desagregation (G1 vs. G2) :

$$
\begin{aligned}
1 = & \frac{\underline{C}_1}{\underline{C}} \times \underline{\gamma}_1 + \frac{\underline{C}_2}{\underline{C}} \times \underline{\gamma}_2 \\
& + \frac{C_{0,1} - \underline{C}_1}{\underline{C}} \times \left(\gamma^{1}_{\text{Fixe\_HB}} - 1\right) + \frac{C_{0,2} - \underline{C}_2}{\underline{C}} \times \left(\gamma^{2}_{\text{Fixe\_HB}} - 1\right) \\
& + \frac{C^{\kappa=1}_{v,1} - C_{0,1}}{\underline{C}} \times \left(\gamma^{\kappa=1}_{v,1} - 1\right) + \frac{C^{\kappa=1}_{v,2} - C_{0,2}}{\underline{C}} \times \left(\gamma^{\kappa=1}_{v,2} - 1\right) \\
& + \frac{C^{\kappa=\kappa_0}_{v,1} - C^{\kappa=1}_{v,1}}{\underline{C}} \times \left(\gamma^{1}_{\text{surco}} - 1\right) + \frac{C^{\kappa=\kappa_0}_{v,2} - C^{\kappa=1}_{v,2}}{\underline{C}} \times \left(\gamma^{2}_{\text{surco}} - 1\right) \\
& + \frac{\underline{C}_1}{\underline{C}} \times \frac{-\Pi_1}{\underline{C}_1} + \frac{\underline{C}_2}{\underline{C}} \times \frac{-\Pi_2}{\underline{C}_2}
\end{aligned}
\tag{11.36}
$$

Service desagregation (EP vs. A) :

$$
\begin{aligned}
1 = & \frac{\underline{C}_{\text{EP}}}{\underline{C}} \times \underline{\gamma}_{\text{EP}} + + \frac{\underline{C}_{\text{A}}}{\underline{C}} \times \underline{\gamma}_{\text{A}} \\
& + \frac{C_{0,\text{EP}} - \underline{C}_{\text{EP}}}{\underline{C}} \times \left(\gamma^{\text{EP}}_{\text{Fixe\_HB}} - 1\right) + \frac{C_{0,\text{A}} - \underline{C}_{\text{A}}}{\underline{C}} \times \left(\gamma^{\text{A}}_{\text{Fixe\_HB}} - 1\right) \\
& + \frac{C^{\kappa=1}_{v,\text{EP}} - C_{0,\text{EP}}}{\underline{C}} \times \left(\gamma^{\kappa=1}_{v,\text{EP}} - 1\right) + \frac{C^{\kappa=1}_{v,\text{A}} - C_{0,\text{A}}}{\underline{C}} \times \left(\gamma^{\kappa=1}_{v,\text{A}} - 1\right) \\
& + \frac{C^{\kappa=\kappa_0}_{v,\text{EP}} - C^{\kappa=1}_{v,\text{EP}}}{\underline{C}} \times \left(\gamma^{\text{EP}}_{\text{surco}} - 1\right) + \frac{C^{\kappa=\kappa_0}_{v,\text{A}} - C^{\kappa=1}_{v,\text{A}}}{\underline{C}} \times \left(\gamma^{\text{A}}_{\text{surco}} - 1\right) \\
& + \frac{\underline{C}_{\text{EP}}}{\underline{C}} \times \frac{-\Pi_{\text{EP}}}{\underline{C}_{\text{EP}}} + \frac{\underline{C}_{\text{A}}}{\underline{C}} \times \frac{-\Pi_{\text{A}}}{\underline{C}_{\text{A}}}
\end{aligned}
\tag{11.37}
$$

Group and Service desagrégation (EP-G1, EP-G2 et EA-G2) :

$$
\begin{aligned}
1 = & \frac{\underline{C}^{1}_{\text{EP}}}{\underline{C}} \times \underline{\gamma}^{1}_{\text{EP}} + \frac{\underline{C}^{2}_{\text{EP}}}{\underline{C}} \times \underline{\gamma}^{2}_{\text{EP}} + \frac{\underline{C}^{2}_{\text{A}}}{\underline{C}} \times \underline{\gamma}^{2}_{\text{A}} \\
& + \frac{C^{1}_{0,\text{EP}} - \underline{C}^{1}_{\text{EP}}}{\underline{C}} \times \left(\gamma^{\text{EP},1}_{\text{Fixe\_HB}} - 1\right) + \frac{C^{2}_{0,\text{EP}} - \underline{C}^{2}_{\text{EP}}}{\underline{C}} \times \left(\gamma^{\text{EP},2}_{\text{Fixe\_HB}} - 1\right) + \frac{C^{2}_{0,\text{A}} - \underline{C}^{2}_{\text{A}}}{\underline{C}} \times \left(\gamma^{\text{A},2}_{\text{Fixe\_HB}} - 1\right) \\
& + \frac{C^{\text{EP},\kappa=1}_{v,1} - C^{1}_{0,\text{EP}}}{\underline{C}} \times \left(\gamma^{\text{EP},\kappa=1}_{v,1} - 1\right) + \frac{C^{\text{EP},\kappa=1}_{v,2} - C^{2}_{0,\text{EP}}}{\underline{C}} \times \left(\gamma^{\text{EP},\kappa=1}_{v,2} - 1\right) + \frac{C^{\text{A},\kappa=1}_{v,2} - C^{2}_{0,\text{A}}}{\underline{C}} \times \left(\gamma^{\text{A},\kappa=1}_{v,2} - 1\right) \\
& + \frac{C^{\text{EP},\kappa=\kappa_0}_{v,1} - C^{\text{EP},\kappa=1}_{v,1}}{\underline{C}} \times \left(\gamma^{\text{EP},1}_{\text{surco}} - 1\right) + \frac{C^{\text{EP},\kappa=\kappa_0}_{v,2} - C^{\text{EP},\kappa=1}_{v,2}}{\underline{C}} \times \left(\gamma^{\text{EP},2}_{\text{surco}} - 1\right) + \frac{C^{\text{A},\kappa=\kappa_0}_{v,2} - C^{\text{A},\kappa=1}_{v,2}}{\underline{C}} \times \left(\gamma^{\text{A},2}_{\text{surco}} - 1\right) \\
& + \frac{\underline{C}^{\text{EP}}_{1}}{\underline{C}} \times \frac{-\Pi^{\text{EP}}_{1}}{\underline{C}_{\text{EP}}} + \frac{\underline{C}^{\text{EP}}_{2}}{\underline{C}} \times \frac{-\Pi^{\text{EP}}_{2}}{\underline{C}_{\text{EP}}} + \frac{\underline{C}^{\text{A}}_{2}}{\underline{C}} \times \frac{-\Pi^{\text{A}}_{2}}{\underline{C}_{\text{A}}}
\end{aligned}
\tag{11.38}
$$

Table 56 : Groups Disaggregation of basic service funding

Weight

	Captive part		Variable part			
	Basic	Captive non basic	PP Economic part	Overcomsumption	Direct support	Total
G1	0.2987	0.0151	0.0835	0.0129	0.2987	
G2	0.7013	0.0193	0.0874	0.0129	0.7013	
All						

Self-financing / Margin rate / Direct support

	Captive part		Variable part			
	Basic	Captive non basic	PP Economic part	Overcomsumption	Direct support	Total
G1	0.552	0.777	1.470	2.218	-0.099	
G2	0.567	1.819	2.225	2.673	0.056	
All						

Contributions (in percentage points)

	Captive part		Variable part			
	Basic	Captive non basic	PP Economic part	Overcomsumption	Direct support	Total
G1	16.5	1.2	12.3	2.9	-2.9	
G2	39.8	3.5	19.5	3.4	3.9	
All						

Table 57 : Services Disaggregation of basic service funding

Weight

	Captive part		Variable part			
	Basic	Captive non basic	PP Economic part	Overcomsumption	Direct support	Total
EP	0.6371	0.0285	0.1440	0.0218	0.6371	
A	0.3629	0.0059	0.0269	0.0040	0.3629	
Ensemble						

Self-financing / Margin rate / Direct support

	Captive part		Variable part			
	Basic	Captive non basic	PP Economic part	Overcomsumption	Direct support	Total
EP	0.557	0.825	1.410	2.073	0.016	
A	0.572	3.934	4.247	4.495	-0.001	
Ensemble						

Contributions (en percentage points)

	Captive part		Variable part			
	Basic	Captive non basic	PP Economic part	Overcomsumption	Direct support	Total
EP	35.5	2.3	20.3	4.5	1.0	
A	20.8	2.3	11.4	1.8	0.0	
Ensemble						

Table 58 : Groups & Services Disaggregation of basic service funding

Weight	EP service						A service				
	Captive part		Variable part				Captive part		Variable part		
	Basic	Captive non basic	PP Economic part	Overcons.	Direct Support		Basic	Captive non basic	PP Economic part	Overcons.	Direct Support
Group 1	0.2987	0.0151	0.0835	0.0129	0.2987						
Group 2	0.3384	0.0134	0.0605	0.0089	0.3384		0.3629	0.0059	0.0269	0.0040	**0.3629**

Self-financing / Margin rate / Direct support

	EP service						A service				
	Captive part		Variable part				Captive part		Variable part		
	Basic	Captive non basic	PP Economic part	Overcons.	Direct Support		Basic	Captive non basic	PP Economic part	Overcons.	Direct Support
Group 1	0.552	0.777	1.470	2.218	-0.099						
Group 2	0.562	0.880	1.326	1.863	0.117		0.572	3.934	4.247	4.495	

Contributions (en percentage points)

	EP service						A service				
	Captive part		Variable part				Captive part		Variable part		
	Basic	Captive non basic	PP Economic part	Overcons.	Direct Support		Basic	Captive non basic	PP Economic part	Overcons.	Direct Support
Group 1	16.5	1.2	12.3	2.9	-2.9						
Group 2	19.0	1.2	8.0	1.7	4.0		20.8	2.3	11.4	1.8	0.0

11.3 Gross subsidies and gross "taxes"

The analysis of the "financing structure" aspects concludes with an examination of the gross subsidies and gross taxes (in fact, the contributions to service funding) generated by the IBT (which is tested/assessed by the user), in relation to the cost to the service of the household subscriber's consumption.

11.3.1 Data processing - reminder

As previously explained, an Incrasing Block Tariff of the social incentive type sets subsidies on, presumably, the access fee and the first consumption blocks, so as to support the affordability of the service, and taxation on the higher consumption blocks in order, on the one hand, to induce large consumers to reduce their consumption and, on the other hand, to balance the service funding (and meet the "water pays for water" principle). In so doing, it generates a system of cross-subsidies with net effects including VAT, the extent of which was measured in section IX. The aim now is to describe and characterize the impacts of this system from the point of view of its gross effects on the service funding (borne by the Operator), i.e. excluding VAT and environmental charges.

For this purpose, it is calculated for each household in the Population module :

(1) the difference between the amount of the subscription (characteristic of the IBT which is tested/assessed by the user) and the amount of the fixed costs per domestic subscriber (which also corresponds to the level of TBSE subscription which would be collected by the operator):

$$c_{i0} = F - \frac{CF}{n} \tag{11.39}$$

This variable is then positive in the (unlikely) case of a tax on the Access Fee and negative in the (more likely) case of a subsidy on the Access Fee. For the purposes of the analysis, the two following truncated variables (in addition to this variable c_{i0}) are calculated:

$$c_{i0}^{+} = \max\left[0, F - \frac{CF}{n}\right] = \begin{cases} F - \dfrac{CF}{n} & if \quad F \geq \dfrac{CF}{n} \\ 0 & if \quad F < \dfrac{CF}{n} \end{cases} \tag{11.40}$$

$$c_{i0}^{-} = -\min\left[0, F - \frac{CF}{n}\right] = \begin{cases} \dfrac{CF}{n} - F & if \quad F \leq \dfrac{CF}{n} \\ 0 & if \quad F > \dfrac{CF}{n} \end{cases} \tag{11.41}$$

The first truncated variable gives the amount of the margin on the Access Fee when there is a charge on the Access Fee and 0 otherwise; the secund truncated variable gives the amount of the subsidy on the Access Fee when there is a subsidy on the Acces Fee and 0 otherwise. It should be noted that these truncated variables are calculated for the EP service, the A service (when households are connected to the sewerage network) and for the consolidated EP / EPA service, with households in group 1 (not connected to the sewerage network) being assigned the values of these variables for the EP service only and households in group 2 (connected to the sewerage network) being assigned the values of these variables for the EPA service.

After breaking down the household's consumption into the sum of its consumption in block 1, its consumption in block 2, its consumption in block 3 ...[49] :

$$q_i^1 = \min\left[q_i, k_1\right] = \begin{cases} q_i & if \quad q_i \le k_1 \\ \\ k_1 & if \quad q_i > k_1 \end{cases}$$

$$q_i^2 = \min\left[k_2 - k_1, \max\left[q_1 - k_1, 0\right]\right] = \begin{cases} 0 & if \quad q_i \le k_1 \\ q_1 - k_1 & if \quad k_1 < q_1 \le k_2 \\ k_2 - k_2 & if \quad q_i > k_2 \end{cases}$$

$$q_i^3 = \min\left[k_3 - k_2, \max\left[q_i - k_2, 0\right]\right] = \begin{cases} 0 & if \quad q_i \le k_2 \\ q_1 - k_2 & if \quad k_2 < q_i \le k_3 \\ k_3 - k_2 & if \quad q_i > k_3 \end{cases}$$

...

the $3 \times p$ following variables are calculated for each household in the Population module:

$$c_{i,j} = \left(\pi_j - c\right) q_i^j \tag{11.42}$$

$$c_{i,j}^+ = \max\left[\left(\pi_j - c\right) q_i^j, 0\right] = \begin{cases} \left(\pi_j - c\right) q_i^j & if \quad \left(\pi_j - c\right) q_i^j > 0 \\ \\ 0 & if \quad \left(\pi_j - c\right) q_i^j \le 0 \end{cases} \tag{11.43}$$

$$c_{i,j}^- = -\min\left[\left(\pi_j - c\right) q_i^j, 0\right] = \begin{cases} \left(c - \pi_j\right) q_i^j & if \quad \left(\pi_j - c\right) q_i^j < 0 \\ \\ 0 & if \quad \left(\pi_j - c\right) q_i^j \ge 0 \end{cases} \tag{11.44}$$

for j varying from 1 to *p*, with *p* the number of consumption blocks of the IBT that is tested/evaluated by the user. To be concrete:

- the variable $c_{i,1} = \left(\pi_1 - c\right) q_i^1$ gives the amount of the subsidy/tax (in fact, the amount of the transfer between the operator and household *i*) which is implemented by the IBT on the consumption of block 1 of household *i*,

[49] It should be noted that, with this formulation, a household located in block 1 has conditional demand in block 2 and above equal to 0, a household located in block 3 has conditional demand in block 3 and above equal to 0, and so on.

- the variable c_{i1}^{+} gives the amount of the margin (and the contribution to service funding) levied by the operator on the consumption of block 1 (equal to 0 when block 1 units are subsidised) of household *i*,
- the variable c_{i1}^{-} gives the amount of the subsidy received by household *i* (and granted by the operator) on its block 1 consumption (equal to 0 when block 1 units are "taxed" (marked up)),

and similarly[50]:

- the variable $c_{i,2} = (\pi_2 - c)q_i^2$ gives the amount of the subsidy/tax (in fact, the amount of the transfer between the operator and household *i*) which is implemented by the IBT on household *i*'s consumption of block 2, the variable c_{i2}^{+} the amount of the margin (and the contribution to service funding) which is levied by the operator on household *i's* consumption of block 2 (equal to 0 when the block 2 units are subsidised and/or when the household's consumption of block 2 is zero), the variable c_{i2}^{-} the amount of the subsidy received by household *i* (and granted by the operator) on its consumption of block 2 (equal to 0 when block 2 units are "taxed" (marked up) and/or when household *i*'s consumption of block 2 is zero),
- the variable $c_{i,3} = (\pi_3 - c)q_i^3$ gives the amount of the subsidy/tax (in fact, the amount of the transfer between the operator and household *i*) which is implemented by the IBT on household *i*'s consumption of block 3, the variable c_{i3}^{+} the amount of the margin (and the contribution to service funding) which is levied by the operator on household *i*'s consumption of block 3 (equal to 0 when the units of block 3 are subsidised and/or when the household's consumption of block 3 is zero), the variable c_{i3}^{-} the amount of the subsidy received by household *i* (and granted by the operator) on its consumption of block 3 (equal to 0 when the block 3 units are "taxed" (marked up) and/or when the household's consumption of block 3 is zero),

etc.

As with Acces Fee, these variables are calculated:

- for the EP service,
- for the A service (when households are connected to the sewerage network),
- for the consolidated service EP / EPA,

with, in the latter case, households in Group 1 (not connected to the sewerage network) being assigned the values of these variables for the EP service only, and households in Group 2 (connected to the sewerage network) being assigned the values of these variables for the EPA service.

Once this information made available (and stored in the Invoices Module), the tool finalises this data processing with:

[50] modulo the fact that consumption in block 2, in block 3 .. by households located in block 1 is zero, consumption in block 3 … by households located in block 2 is zero, and so on.

(1) calculation of the total (gross) subsidies received by household *i* on its consumption :

$$s_q = c_{i1}^- + c_{i2}^- + c_{i3}^- + \ldots = \sum_{j=1}^{p} c_{ij}^- \quad (11.45)$$

also designed as Subsidy "DAE" (for "Droit d'Accès Exclu");

(2) calculation of the total (gross) subsidies received by household *i* on the Acces Fee and its consumption:

$$s = c_{i0}^- + c_{i1}^- + c_{i2}^- + c_{i3}^- + \ldots = c_{i0}^- + s_q \quad (11.46)$$

also designed as Subsidy "DAI" (for "Droit d'Accès Inclu");

(3) the total of "taxes" (in fact, gross contributions to service funding) paid on its consumption by household *i* :

$$t_q = c_{i1}^+ + c_{i2}^+ + c_{i3}^+ + \ldots = \sum_{j=1}^{p} c_{ij}^+ \quad (11.47)$$

also designed as Contribution (to service funding) "DAE" (for "Droit d'Accès Exclu") ;

(4) the total of "taxes" (in fact, gross contributions to the service funding) paid by household *i* for the Acces Fee and its consumption:

$$t = c_{i0}^+ + c_{i1}^+ + c_{i2}^+ + c_{i3}^+ + \ldots = c_{i0}^+ + t_q \quad (11.48)$$

also designed as Contribution (to service funding) "DAI" (for "Droit d'Accès Inclu") ;

for the EP service, the A service and the EP/EPA service (in turn).

It should be noted that, in terms of data presentation, the value of the variable c_{i0} is most often shown, which should then be interpreted as giving the amount of the subsidy on the Acces Fee when negative (the most likely case), and the amount of the tax (margin) on the Access Fee when positive. Besides, the values in question here are exclusive of taxes and environmental charges, so that focus is on the system of subsidies/taxations carried out by the Operator with the definition of its pricing policy (for the EP service and A service). However, this treatment is also applied to the amounts of VAT (which are collected by the Operator for the benefit of the State), because the VAT mechanism implements also a system of subsidies/taxations on State's side, with the introduction of the IBT, but not for the environmental charges (levied by the Water Agency) because excise duty mechanism means that the Water Agency's contribution to the system of subsidies/taxes generated by the Operator (with the introduction of the IBT) is (identically) zero. Finally, in order to facilitate reading, the data on gross subsidies can be presented in negative (this convention because it constitutes an outgoing cash flow for the Operator).

11.3.2 Initial Findings– Descriptive Statistics

Once the data has been processed, the tool (as for the other fields of analysis) begins by providing basic descriptive statistics for the general EP / EPA service, with :

(1) on the **gross subsidies** side :

- the percentage of beneficiaries on Acces Fee, the percentage of "DAE" beneficiaries, the percentage of "DAI" beneficiaries,
- the average of the (gross) subsidies granted on the Access Fee, the average of the (gross) subsidies granted on consumption, the average of the subsidies granted in fine to the household (adding the subisidies granted on the Acces Fee and on the consumption) …
- the variance of subsidies granted on the Access Fee …

etc. (see Table 59, page 214; the 3 variables of interest are the amount of the subsidy on the Access Fee, the amount of the subsidy on consumption (or "DAE") and the amount of the subsidy "DAI" that is ultimately paid to the household) and,

(2) concerning the **gross margin** side :

- the percentage of contributors on Access Fee, the percentage of "DAE" contributors, the percentage of "DAI" contributors,
- the average of the "taxes" (margins) that are levied on the Access Fee, on consumption ("DAE"), on the household ("DAI") ...
- the variance of the "taxes" (margins) that are levied (realised) on the Access Fee …

etc. (see Table 60, page 215; the 3 variables of interest are the amount of the tax (margin) on the Access Fee, the amount of the tax (margin) on consumption (or "DAE") and the amount of the tax (margin) which is ultimately levied on the household (or "DAI"). Besides,

(3) these figures are calculated:

- for the population of the domestic customer as a whole

and also (where it makes sense):

- on the sub-population of beneficiaries alone for the Subsidy component,
- on the sub-population of taxpayers alone for the "Taxation" component.

The following points should be borne in mind.

(1) Firstly, because these statistics are calculated (initially) for the general EP / EPA service, not all households benefit from the same treatment with regard to the Access Fee.

In particular, and in the (most likely) case where the EP tariff and the A tariff both subsidise the Access Fee, a household connected to the sewerage network (member of Group 2) benefits from a higher subsidy on the Access Fee compared to the one granted to a household that is not connected to the sewerage network (member of Group 1). For this reason, (i) the average calculated and displayed by the tool differs from the amounts of subsidies paid on the Access Fee by the EP tariff and the EPA tariff with, therefore, (ii) a standard deviation (and a Gini index) that is not equal to 0.

(2) Secondly, because one deals here with gross subsidies, the percentage of DAI beneficiaries is automatically equal to the percentage of beneficiaries on the Access Fee, with a 100% value when EP tariff and A tariff subsidise both the Access Fee. Similarly, the percentage of DAE beneficiaries is automatically equal to 100% when EP tariff and A tariff both subsidise some consumption units.

In practice, however, it may happen that the tariff implemented by the operator subsidises the Access Fee and taxes all the units consumed, starting with block 1. In such a case, the tool will display the values 100, 0 and 100 for, successively, the percentage of beneficiaries on the Access Fee, the percentage of DAE beneficiaries and the percentage of DAI beneficiaries (100% of households are subsidised on the Access Fee, 0% of households are subsidised on their consumption and 100% of households receive some subsidies). Besides, and even though they are unlikely, the operator may charge the Access Fee and subsidising all consumption units. In such cases, the tool will display the values 0 for the percentage of beneficiaries on the Access Fee, 100 for the percentage of DAE beneficiaries and 100 for the percentage of DAI beneficiaries.

(3) While the gross taxation/margin on the Access Fee mirrors the gross subsidy on the Acces Fee, this no longer holds when it comes to consumption. Indeed, the financial equilibrium requires subsidies to be self-financing through "taxes", and there are households that will both receive subsidies (on the Access Fee and/or the first consumption units) and be taxed on their consumption. At the same time, calibration of the tariff system may also mean that small consumers whose consumption are located in the first subsidised consumption blocks are not "taxed" on their consumption. Given these elements, the percentage of DAE contributors is positive, is not equal to 100% (except in some special cases) and, in all cases, is not equal to 100 minus the percentage of DAE beneficiaries (unlike in the case of net subsidies and net taxation).

(4) In view of this last point, it matters to calculate and display statistics on the population of households that actually contribute to service funding, in addition to those calculated for the population (and that includes households that are fully subsidised on their consumption). For this reason, the tool displays descriptive statistics on gross contributions for the sole sub-population of "effective" contributors (see the last three columns of Table 60).

(5) Insofar as focus is on gross subsidies, all households are automatically subsidised as soon as the EP tariff and the A tariff both subsidise the Acces Fee and/or consumption (with the setting of tariff parameters for the first consumption blocks). Accoridngly, the tool does not spontaneously display some statistics calculated for the sole sub-population of beneficiaries, because, in this case, the latter merges simply with the general population and the information would appear to be redundant (values would be identical to those displayed, right next, for the population as a whole). However, in order to deal with atypical cases in which, for example, the EP tariff is subsidised everywhere and the A tariff is taxed everywhere (with a wastewater service financing the drinking water service), the user can display a complete table (with statistics calculated for the population as a whole and for the sole sub-population of beneficiaries) by clicking on a button.

Once these initial elements (on the general service EP / EPA) introduced, the tool completes this (aggregated) information (in the general population) by displaying all these descriptive statistics for (i) the drinking water service, (ii) the wastewater service, (iii) the drinking water and wastewater service, and a contingency table (flow matrix) cross-referencing subsidies/taxations and services (EP vs. / A). The latter concludes this "Initial Findings" section.

	Access Fee	"DAE"	"DAI"
% of beneficiaries	100.00	100.00	100.00
Mean	48.39	0.33	48.72
Median	28.33	0.33	28.66
Min	28.33	0.19	28.58
Max	72.08	0.33	72.41
Q1	28.33	0.33	28.66
Q3	72.08	0.33	72.41
D1	28.33	0.33	28.66
D9	72.08	0.33	72.41
F (Mean)	54.1	1.4	54.1
Variance	475.0812	0.0001	475.0648
Standard deviation	21.80	0.01	21.80
MAPE	21.72	0.00	21.72
Coeff of Variation	0.450	0.024	0.447
Interquartile range	43.74	0.00	43.74
Interdecile range	43.74	0.00	43.74
Yule coefficient	43.74	#DIV/0!	43.74
Gini index			
Schutz coefficient	22.4	0.2	22.3
Interdecile ratio	2.54	1.00	2.53
Interdecime ratio			
S80 / S20 ratio			

Table 59 : Operator gross subsidies -- General Service "EP/EPA" -- Main Descriptive Statistics

	Total Population			Contributors Population		
	Access Fee	"DAE"	"DAI"	Access fee	"DAE"	"DAI"
% of contributors	0.00	99.78	99.78	**	**	**
Mean	0.00	48.70	48.70	n.a.	48.81	48.81
Median	0.00	42.83	0.00			
Min	0.00	0.00	0.00			
Max	0.00	236.48	236.48			
Q1	0.00	22.11	22.11			
Q3	0.00	63.46	63.46			
D1	0.00	12.85	12.85			
D9	0.00	92.18	92.18			
F (Mean)	n.a.	n.a.	n.a.			
Variance	0.0000	1208.7309	1208.7309			
Standard deviation	0.00	34.77	34.77			
MAPE	0.00	25.94	25.94			
Coeff of Variation	n.a.	0.714	0.714			
Interquartile range	0.00	41.35	41.35			
Interdecile range	0.00	79.33	79.33			
Yule coefficient	#DIV/0!	20.13	64.00			
Gini index						
Schutz coefficient	0	26.6	26.6			
Interdecile ratio	Undefined	7.17	7.17			
Interdecime ratio						
S80 / S20 ratio						

Table 60 : Operator gross margins -- General EP/EPA services -- Main Descriptive statistics

	"EP"	"A"	
Subvention	1’371’493.55	1’106’711.5	2’478’205.05
Taxation	1’325’223.76	1’108’119.06	24’33’342.82
	2’696’717.31	2’214’830.56	4’911’547.87

Table 61 : Subsidies/taxes by service (contingency table)

11.3.3 Who? What? How?

The second stage of the analysis consists of completing the previous initial elements by producing tables and calculating statistics organised according to a "Who ? What? How?" Interpretative Framework. **The first question** ("Who?") aims to determine:

- Which households benefit from gross subsidies (presumably all) and (above all) for what amounts?

and also:

- Which households are "taxed" (margined) and for what amounts?

To this end, the tool displays information on :

- the percentages of beneficiaries, the averages and the variances of the subsidies Access Fee Excluded and Acces Fee Included,
- the percentages of contributors, the averages and the variances of the margins, Access Fee Excluded and Acces Fee Included

by breaking down the population (of domestic subscribers) by customer segment (G1 vs. G2), by standard of living (Poor / Non-Poor) and by customer segment and standard of living (G1-Poor, G1-Non Poor, G2-Poor, G2-Non Poor). The tool also displays decompositions of the related averages and variances.

The second question ("What") aims at identifying what is subsidised and what is 'taxed' (marked up). To this end, the tool calculates gross subsidies and gross taxes for each of the households in the Population Module on :

(1) Access Fee,

(2) basic consumption $\underline{q}_i$,

(3) captive but non basic part of water consumption $q_{0,i} - \underline{q}_i$,

(4) variable component of consumption excluding overconsumption $q_i^{\kappa=1} - q_{0,i}$,

(5) over-consumption (due to tariff misperception) $q_i^{\kappa=\kappa_0} - q_i^{\kappa=1}$.

The sum of the subsidies (mutatis mutandis of the margins) on the Access Fee and basic consumption then gives the total gross subsidies granted to the basic service, while the sum of the subsidies (mutatis mutandis of the margins) on the captive but non-basic part of the demand (by default, the uses of water linked to garden maintenance and swimming pool maintenance) and the variable part of the demand $q_i^{\kappa=\kappa_0} - q_0^i$ gives the total subsidies granted to the non-basic service.

Table 62 : Breakdown of Operator gross subsidies and gross margins for general service "EP/EPA": basic uses vs. other uses (comfort and luxury)

	Access Fee		Basic consumption		Basic service		Non basis onsumption		Total	
	Mean	Total	Mean	Total	Mean	Total	Mean	Total	Mean	Total
Subsidies	48.39	2315410.70	0.30	14234.01	48.69	2329644.72	0.03	1522.52	48.72	2331167.23
"Taxes"	0.00	0.00	8.89	425294.53	8.89	425294.53	39.81	1904934.54	48.70	2330229.07
									-0.02	-938.16

Table 63 : Leakage rates, PPV, FOR et NPV

Subsidies	"DAI"	"DAE"	Taxes	"DAI"	"DAE"
Leakage rate	0.07	9.7	FOR	18.25	18.25
PPV	99.93	90.3	NPV	81.75	81.75
	100.00	100.00		100.00	100.00

Table 64 : Breakdown of average subsidies (basic service vs. non-basic consumption)

	Basic service			Access Fee			Basic consumption		
	Mean	%	Effective mean	Mean	%	Effective mean	Mean	%	Effective Mean
Suisidies	48.69	100.00	48.69	48.39	100.00	48.39	0.30	100.00	0.30
Margins	8.89	75.55	11.77	0.00	0.00	#DIV/0!	8.89	75.55	11.77
	Non Basic consumption								
	Mean	%	Effective mean						
Subsidies	0.03	44.10	0.07						
Margins	39.81	99.78	39.90						
	Total consumption								
	Mean	%	Effective mean						
Subsidies	0.33	100.00	0.33						
Margins	48.70	99.78	48.81						

Using this information (stored in the Invoices module), the tool first looks at subsidies and taxes on the basic service. To this end, mean, variance and Gini index of the following variables of interest are displayed :

- the subsidy on the Access Fee (which enters in the definition of the basic service),
- the (gross) subsidy Access Fee Excluded, that is the amount of gross subsidies paid on basic consumption,
- the (gross) subsidy Access Fee Included, that isi the amount of support for the provision of basic service of which households benefit,

as well as :

- the (potential) margin on the Access Fee (probably equal to 0)
- the gross margin levied on cubic metres for basic uses, i.e. the amount of contributions to service funding levied on the basic consumption (this variable accounts for exclusion errors in value);
- the amount of contributions to service funding collected from the marketing of the provision of basic service by the operator (equal to the previous value, provided the Access Fee is subsidised) ;

for, successively, the population of Household as a whole and the sub-population of contributors. Next, all this processing is reproduced for non-basic consumption, and thereafter the tool proceeds to:

- average decomposition,
- decomposition by sources (factors) of variance,
- decomposition by sources (factors) of Gini index (also known as Rao decomposition)

of the gross subsidies and of gross margins by services (EP service vs. A service).

The last item relates to the "How" issue which is addressed with a contingency table showing the distribution of subsidies and taxes on Acces Fee, basic consumption, basic service (consolidation of the first two items), and non-basic consumption (see Table 62, page 217). On this basis, the tool calculates several indicators (similar to the ones applied in Section 8.3.2 to assess the proper calibration of the tariff), including leakage rates and FORs (see Table 63 on next page), and then breaks down the apparent averages to show "effective" averages (see Table 64 on next page).

11.3.4 "Good taxes" vs. "Bad" subsidies

The tool concludes this analysis with a "Who, What **and** How" table providing a summary view of good and bad subsidies/taxations, and also showing support for basic services and direct support.

To this end, it is first displayed a table showing the mass of flows with households classified by consumption deciles (see Table 65 on next page). Next, these data are reprocessed to feed a second table in which :

(i) values are expressed as a percentage of total subsidies/taxes (adjusted for the value of operating income),

	T_q^+	T_q^-	T_q	S_q^+	S_q^-	S_q	$T_q + S_q$	C_{i0}	$C_{i0} + T_q + S_q$	**Direct Sipport**	**Final**
d1	80552.33	50469.41	131021.74	-1487.21	-75.44	-1562.65	129459.09	-227563.23	-98104.14	71,80	-98032.34
d2	122861.79	54979.98	177841.77	-1509.10	-72.53	-1581.63	176260.14	-250412.55	-74152.41	80,30	-74072.11
d3	163531.35	50788.22	214319.58	-1482.68	-103.16	-1585.85	212733.73	-232133.10	-19399.37	90,59	-19308.77
d4	165557.71	54270.75	219828.46	-1487.46	-98.39	-1585.85	218242.62	-245842.68	-27600.07	88,31	-27511.76
d5	193476.76	50677.49	244154.25	-1489.03	-96.82	-1585.85	242568.41	-227563.23	15005.17	97,47	15102.65
d6	200913.10	45200.15	246113.25	-1406.97	-177.20	-1584.17	244529.09	-250412.55	-5883.46	92,37	-5791.09
d7	205872.88	40123.20	245996.08	-1362.33	-219.63	-1581.96	244414.12	-241272.82	3141.30	93,08	3234.39
d8	217434.60	29856.95	247291.56	-1410.19	-175.66	-1585.85	245705.71	-218423.51	27282.20	102,29	27384.49
d9	236983.62	25569.86	262553.48	-1317.23	-268.62	-1585.85	260967.63	-218423.51	42544.12	104,19	42648.31
d10	317750.39	23358.51	341108.90	-1281.82	-235.08	-1516.90	339592.00	-203363.52	136228.48	117,75	136346.23
Total	1904934.54	425294.53	2330229.07	-14234.01	-1522.52	-15756.53	2314472.54	-2315410.70	-938.16	938,16	0.00

Table 65 : "Good" and "bad" subsidies / contributions to service funding

	Access Fee			Basic units	Basic service			Non basis service				
	C_{i0}	S_q^+	T_q^-	$S_q^+ + T_q^-$	$C_{i0} + S_q^+ + T_q^-$	S_q^-	T_q^+	$S_q^- + T_q^+$	C_{iq}	$C_{i0} + C_{iq}$	**Direct support**	**Final**
d1	-9.8	-0.0638	2.1650	2.1	-7.7	0.0	3.5	3.5	5.6	-4.21	0.0031	-4.2
d2	-10.7	-0.0647	2.3585	2.3	-8.4	0.0	5.3	5.3	7.6	-3.18	0.0034	-3.2
d3	-10.0	-0.0636	2.1787	2.1	-7.8	0.0	7.0	7.0	9.1	-0.83	0.0039	-0.8
d4	-10.5	-0.0638	2.3281	2.3	-8.3	0.0	7.1	7.1	9.4	-1.18	0.0038	-1.2
d5	-9.8	-0.0639	2.1739	2.1	-7.7	0.0	8.3	8.3	10.4	0.64	0.0042	0.6
d6	-10.7	-0.0604	1.9389	1.9	-8.9	0.0	8.6	8.6	10.5	-0.25	0.0040	-0.2
d7	-10.3	-0.0584	1.7212	1.7	-8.7	0.0	8.8	8.8	10.5	0.13	0.0040	0.1
d8	-9.4	-0.0605	1.2808	1.2	-8.1	0.0	9.3	9.3	10.5	1.17	0.0044	1.2
d9	-9.4	-0.0565	1.0969	1.0	-8.3	0.0	10.2	10.2	11.2	1.83	0.0045	1.8
d10	-8.7	-0.0550	1.0020	0.9	-7.8	0.0	13.6	13.6	14.6	5.84	0.0051	5.8
Total	-99.3	-0.6106	18.2438	17.6	-81.7	-0.1	81.7	81.7	99.3	-0.04	0.0402	0.0

Table 66 : "Good" and "bad" subsidies / contributions to service funding (as a percentage of subsidies/taxes adjusted for the value of operating income)

(ii) households are ranked by standard of living deciles

(iii) flows are reclassified so as to show the amount and composition of support for the basic service (by non-basic uses) generated by the IBT which is assessed/tested by the user.

On this last point , see Table 66, page 220. The tool concludes by calculating the same indicators as those used to assess inclusion and exclusion errors in volume for confusion matrices:

	Subvention A	Taxation A	
Basic Service	$S_{\underline{q}} - \Pi^-$	$T_{\underline{q}}$	
Other uses	$S_{q-\underline{q}}$	$T_{q-\underline{q}} - \Pi^+$	
	$S_q - \Pi^-$	$T_q - \Pi^+$	$S_q + T_q - \left(\Pi^- + \Pi^+\right)$

Table 67 : Inclusion and exclusion errors in value - Access Fee Excluded

	Subvention A	Taxation A	
Basic Service	$C_0^- + S_{\underline{q}} - \Pi^-$	$C_0^+ + T_{\underline{q}}$	
Other uses	$S_{q-\underline{q}}$	$T_{q-\underline{q}} - \Pi^+$	
	$S - \Pi^-$	$T - \Pi^+$	$S + T - \left(\Pi^- + \Pi^+\right)$

Table 68 : Inclusion and exclusion errors in value - Access Fee Included

with $\Pi^- = -\min\left[\Pi, 0\right]$ and $\Pi^+ = \max\left[\Pi, 0\right]$ to adjust downwards the "right" subsidies in the case of a deficit, the "right" taxes in the case of a surplus.

XII –SCALING UP

12.1 Motivations

As it stands, the use of the MMS should allow (with adaptations) an improvement in local management, at the level of the (French) municipalities and inter-municipalities which are currently in charge of the pricing of the drinking water and wastewater services[51]. However, with the implementation of the project, it was clear that the tool did not fully meet the needs of some stakeholders who expect more macroscopic information on the effectiveness of the various pricing policies that can be implemented on the scale of the territory/catchment area.

[51] It is to note that, for the French case, the law NOTRe ("**N**ouvelle **O**rganisation **T**erritoriale de la **R**épublique" in French, New Territorial Organisation of the Republic) of 2015 provided for the mandatory transfer of water and sanitation responsibility to communities of municipalities and urban areas on 1 January 2020. A new law, pending enactment, ends this obligation for municipalities that have not yet transferred their authority, with no possibility of reversal for municipalities that have already transferred their authority (the text in question was adopted at second reading by the Senate on 2 April 2025).

With regard to the study site of Reunion Island specifically[52], some stakeholders involved in integrated water resource management come in mind, foremost among which is:

(1) the local Water Agency (Office de l'Eau de La Réunion), which collects and redistributes eco-taxes (excise duties for the protection of aquatic environments, aid for investment in networks and pollution reduction),

but also:

(2) certain decentralised government departments such as the Regional Health Agency (ARS) or the Department of the Environment, Planning and Housing (DEAL)),

and:

(3) certain local authorities (Regional Council, Departmental Council) which can implement specific social policies with assistance programmes (water vouchers, assistance with unpaid bills).

The latter could be initiated (and supervised) by national regulations (like the FSL (Fond Solidarité Logement, Housing Solidarity Fund)). At the same time, this issue of evaluating the performance of a set of pricing policies in a specific geographical area is also relevant for areas smaller than river basins. This applies in particular (4) to Local Water Commissions, but also and above all, (5) to inter-municipal bodies, which are now in charge of the management of the service (see above). Within these latter structures, different municipal rates continue to apply, with tariff equalisation expected in the long term (which is not without its difficulties given that service costs can vary greatly from one municipality to another).At first glance, the scaling up of the MMS should enable all these stakeholders to obtain as complete a picture as possible of the socio-economic performance of the various pricing policies implemented in Reunion Island and, based on this information, to better target their support measures.

To meet this need, it is here to capitalize on and scale up the micro-simulation model with:

(1) the econometric estimation of the water demand functions of the households living on Reunion Island (and not only for one city of the Island), insofar as Reunion Island is regarded as a complete watershed.

This task, planned as part of the project, was completed using the latest available data provided by the national surveys. The econometric model combines (i) a specific Tariffs database containing the IBTs for the 24 municipalities of Reunion Island, with (ii) a Household database (hosted on the CASD's secure platform) which provides information on the socio-economic characteristics of respondents and the geographical location of households (matching with the Tariffs database then makes it possible to identify the IBT they face) and (iii)) a global demand equation which, in addition to the usual variables, includes a set of `Municipality' dummies that account for a certain heterogeneity in demand behaviour (technically, while having a common structure, each city is assigned a specific demand function).

[52] Reunion Island is a separate watershed. It also has a Regional Council, a Departmental Council (both of which have now responsibilities in the field of water), a Water Agency, a Basin (and Biodiversity) committee, several decentralised State services with specific responsibilities in the field of water (including the DEAL, linked to the Ministry of the Environment, and the ARS, linked to the Ministry of Health), five inter-municipal communities and 24 municipalities. The island of La Réunion is also one of the outermost regions (ORs) of the European Union.

(2) the setting of an aggregated dashboard, fed by local dashboards, and performance indicators computed at the territorial level of the Reunion Island area and, above all, broken down (where possible) into its constituent elements on the basis of spatial (geographical) distribution.

With the econometric demand model for Reunion Island (1), it is possible to feed the dashboard that measures the socio-economic performance of a pricing policy implemented by a municipality for the Reunion Island as a whole. This provides an aggregated dashboard and, through it, a multi-dimensional measure of the performance of all pricing policies that are implemented by all municipalities across the territory. This aggregated dashboard may be used to feed into the socio-economic sections of some written documents part of the local SDAGE (Water Development and Management Master Plan) and, using a dedicated geo-visualisation tool, measure potential disparities and spatial inequalities (in terms of public drinking water and sanitation service pricing) across the entire territory of Reunion Island.

Once this point stated, it is understood that, for the purposes of analysis and also the design of public policy, it matters to link the overall performance of all pricing policies at the level of Reunion Island with the performance of each pricing policy implemented locally or, in short, to link the aggregate scoreboard to the disaggregated scoreboards feed at the level of each city. In concrete terms, and focusing on unaffordability (for instance), the aim is:

(1) to calculate the percentage of households facing an affordability issue across the whole of Reunion Island, that is:.

$$H_{\text{PAR}} = \frac{n_{\text{PAR}}}{n}$$

with n_{PAR} the number of households residing in Reunion island and facing an affordability issue as defined by the PAR, i.e. the number of households spending more than 3% (for example) of their income to meet their basic water needs,

and, insofar as water tariffs and the socio-economic composition of the population differ from one city to another,

(2) link this aggregate figure, obtained for Reunion Island as a whole, to the 24 household headcount ratios calculated for each municipality, that is:

$$H^1_{\text{PAR}} = \frac{n^1_{\text{PAR}}}{n_1},\ H^2_{\text{PAR}} = \frac{n^2_{\text{PAR}}}{n_2},\ \ldots\ H^{24}_{\text{PAR}} = \frac{n^{24}_{\text{PAR}}}{n_{24}}$$

with:

- n_1 the number of household customers in City 1 (Commune (Les) Avirons, if we use alphabetical order) who apply a specific IBT given by the tariff function $T_1(q)$, n_2 the number of household customers in City 2 (Bras Panon) who apply a specific IBT given by the tariff function $T_2(q)$, …, n_{24} the number of household customers in City 24 (Trois Bassins) who apply a specific IBT given by the tariff function $T_{24}(q)$;

- n^1_{PAR} the number of households in City 1 facing an affordability issue with the tariff $T_1(q)$ of City 1, n^2_{PAR} the number of households in City 2 facing an affordability issue with the tariff $T_2(q)$ of

City 2, …, n_{PAR}^{24} the number of households in City 24 facing an affordability issue with the tariff $T_{24}(q)$ of City 24

- H_{PAR}^{1} , H_{PAR}^{2} , …, H_{PAR}^{24} the water unaffordability rates, calculated for the household customers in City 1, in City 2, … up to City 24.

This question is answered in the affirmative because we have:

$$H_{\text{PAR}} = \frac{n_{\text{PAR}}}{n} = \frac{n_{\text{PAR}}^{1} + n_{\text{PAR}}^{2} + \ldots + n_{\text{PAR}}^{24}}{n_1 + n_2 + \ldots + n_{24}} = \frac{n_1}{n}\frac{n_{\text{PAR}}^{1}}{n_1} + \frac{n_2}{n}\frac{n_{\text{PAR}}^{2}}{n_2} + \ldots + \frac{n_{24}}{n}\frac{n_{\text{PAR}}^{24}}{n_{24}}$$
$$= f_1 \times H_{\text{PAR}}^{1} + f_2 \times H_{\text{PAR}}^{2} + \ldots + f_{24} \times H_{\text{PAR}}^{24}$$

i.e. the water unaffordability rate of Reunion Island is a weighted average of the water unaffordability rates of the municipalities in Reunion Island, with the weighting coefficient being the weight of the municipality in the distribution of the household population across all municipalities on the Island. Highlighting this breakdown provides useful information for bodies operating at a centralised level, as it enables them to:

(1) identify the municipalities with the highest level of water unaffordability as defined by the PAR, i.e. where needs are greatest

and, also, with the calculation of contributions:

(2) allocate (limited) resources efficiently, with a view to maximising the impact on the aggregate water unaffordability rate, which is in itself a key performance indicator that can be used to measure the effectiveness of a public policy such as, for example, the roll-out of a programme of assistance with water bill payments that can be targeted at poor households, but also programmes to support investment in networks (that are managed by the local water agency) which, through their impact on the local cost of the service, also have an impact on the price of water and, ultimately, on the local and global affordability of the water tariff system.

The point is that this approach is not limited to the percentage of households facing an affordability issue. It also applies, for example, to (1) the affordability deficit, which, because it is an average, will verify:

$$\overline{e}_{\text{PAR}} = \frac{n_1}{n} \times \overline{e}_{\text{PAR}}^{1} + \frac{n_2}{n} \times \overline{e}_{\text{PAR}}^{2} + \ldots + \frac{n_{24}}{n} \times \overline{e}_{\text{PAR}}^{24}$$

i.e. the affordability deficit in Réunion Island, $\overline{e}_{\text{PAR}}$, is a weighted average of the affordability deficits of the 24 municipalities on the Island $\overline{e}_{\text{PAR}}^{1}$, $\overline{e}_{\text{PAR}}^{2}$, …, $\overline{e}_{\text{PAR}}^{24}$, using a weighting system identical to the Household Headcount Ratio and, with an adjustment to the weighting system, for (2) the effective affordability deficit, which breaks down as follows:

$$\overline{e}_{\text{PAR}}^{eff} = \frac{n_{\text{PAR}}^{1}}{n_{\text{PAR}}} \times \overline{e}_{\text{PAR}}^{eff,1} + \frac{n_{\text{PAR}}^{2}}{n_{\text{PAR}}} \times \overline{e}_{\text{PAR}}^{eff,2} + \ldots + \frac{n_{\text{PAR}}^{24}}{n_{\text{PAR}}} \times \overline{e}_{\text{PAR}}^{eff,24}$$

i.e. the effective affordability deficit of Reunion Island, $\overline{e}_{\text{PAR}}^{eff}$, is a weighted average of the effective affordability deficits of the 24 municipalities of Reunion Island $\overline{e}_{\text{PAR}}^{eff,1}$, $\overline{e}_{\text{PAR}}^{eff,2}$, …, $\overline{e}_{\text{PAR}}^{eff,24}$, but with weightings that are now given by the weight of the municipality in the population of households facing an affordability issue. Furthermore, these breakdowns, insofar as they relate to proportions and averages, also apply to other points of performance, such as for instance the incentive effect of the tariff system, with a study of the impacts on consumption linked to the implementation of IBTs, compared to TBSEs, and the issue of overconsumption (that come with) and for which awareness-raising programmes (or nudging campaigns) may be deployed in light of the information provided (and that can help in the design of these non-tariff instruments).

With this programme in place, it should be borne in mind that (i) not all of the indicators in the aggregate dashboard can be broken down (such as, for instance, the median of affordability gap), (2) some of the measures used in the dashboard (such as variance or the Gini index) can be broken down, but with a slightly more complex disaggregation, and (3) the scaling up operation should not be limited to feeding data into an aggregate dashboard and breaking down some indicators that are already known to be broken down. Thus, in addition to points (1) and (2) described above, scaling up will also require:

(3) the use of new specific indicators (used in spatial economics and, more generally, in geography) to measure potential disparities and spatial inequalities across the whole territory of Reunion Island (i.e. at the level of the river basin), for some or even all of the five previous performance points related to the European Water Framework Directive.

Indeed, the scaling up of the micro-simulation model will take place at the basin level, in a framework where (i) the issue of territorial inequalities will constitute a matter of interest with, at the same time, (ii) a large number of stakeholders / decision-makers who act at different levels in the institutional system that organizes water policy,

This last dimension calls for reflection on the architecture of the tool, that must be consistent with the multi-level multi-stakeholder governance system that organises water policy, and within which a large number of decision-makers act at their own level of intervention in specific areas. Ideally, this need will have to be met by offering the possibility to the user to select a geographical area of assessment and a spatial scale of breakdown (for some the indicators making up the dashboard) as specified in the table below[53].

Analysis level / Spatial breakdown	Basin	Intercommunality	City
Basin	√	√	√
Intercommunality		√	√
City			√

Table : The "almost" ideal spatial division

This spatial division is based on the decision-making system of Reunion Island, of the French system more generally, with (i) tariff policies that are designed at the level of cities and intermunicipalities and (ii) some non-tariff measures that the up-scaled MSM aims also to inform. This includes (i) the subsisides for network improvement that affect production cost, and therefore ultimately the price of water, but also (ii) the aids for unpaid bills that are implemented by the General Council, with a funding mechanism based solely on voluntary contributions from the authorities responsible for organising the drinking water and wastewater services that

[53] In addition to the administrative organisation, there is also a specific typology for Reunion Island, with four micro-regions (North, South, East and West) which reflect real economic differences and often form the basis for local public debate. The latter could be extended to five micro-regions, taking into account the specific nature of the "Les Hauts" area, which is similar to the urban vs. rural distinction in mainland France.

questions. Finally, there is a specific issue for Reunion Island and that relate to (iii) the effects of the (potential) standardization of water and sanitation tariffs within the basin (several local public decision-makers are campaigning, in the name of inter-municipal solidarity, for the setting of a single water and wastewater tariff across the island, in a context where service costs can vary greatly from one municipality to another).

12.2 Constraints and challenges for the upscaling operation

The micro-simulation tool for one municipality is very thorough and provides a lot of fine-grained information of the effects that the water and sanitation tariff has on a municipality's population. However the goal for this upscaling operation is to switch the focus from a municipality to a wider entity such as a water agency. In order to better understand the architecture design choices, it is important to review key differences between the base simulation model and the intended upscaled model.

First of all, we are not interested in one tariff anymore, but multiple tariffs. This can go from 24 tariffs if we are interested in La Réunion, to more than 35 000 tariffs if we are looking at Metropolitan France. This of course dramatically increases the computational power needed and smart design choices have to be made to limit that increase. Another aspect that increases the complexity is that a tariff can be included multiple times in an aggregation. For example, the municipality of Saint-Denis, La Réunion, is included both in the aggregation of the whole island, and in the aggregation per arrondissement. It is quite obvious that it is not a smart choice to recompute the results of a single simulation everytime it is included in a new aggregation.

We have mentioned aggregation a few times already and it would be now a good time to explain the scope of the aggregation for our case of interest : France. In order to understand how is France managed, we can refer to the official internet platform "Vie publique" which provides official information on the state organization :

"The French administrative system is structured around three standard territorial divisions: the arrondissement, the department, and the region. Other divisions are referred to as specialized divisions due to their specific purpose (such as water management, for example)."[54]

We can also refer to the INSEE definition of what is an arrondissement[55], which adds that it is composed of municipalities (since 2015). This provides clear indication on the official administrative territories structure that we are interested in within the scope of water tariff management.

It is worth mentioning the bottom division "IRIS INSEE division" which is a division that can be defined as a "fundamental unit for dissemination of infra-municipal data"[56] by INSEE and is widely used in socio-economic studies. The structure shown in the tree is the strict hierarchy that needs to be used within our scope. However, it is important to remember that other custom defined divisions could be used such as the inter municipalities divisions which are groups of municipalities. By knowing the French administrative structure, we can now explain what we mean by aggregation more clearly. Since we know that at the moment the tariffs are defined at the municipality scale, we can use it as an example. If we imagine three municipalities named A,

[54] https://www.vie-publique.fr/fiches/20231-quelles-sont-les-circonscriptions-administratives
[55] https://www.insee.fr/fr/metadonnees/definition/c1912
[56] https://www.insee.fr/en/metadonnees/definition/c1523

B and C that are part of an arrondissement D, the base simulation model is able to compute the results A, B and C, but not the results D since it does not have a tariff and is composed of municipalities. In this case, by aggregation we mean that we take the results A, B and C and use their weighted contribution to the results D.

Now that the constraints are well known, it is time to explain how they led to the design choices that were made.However, before we move further it important to explain that this kind of work has never been done before and is therefore exploratory. In order to reduce the architecture design complexity, it has been decided that we would focus on the first indicator which is the abordability. The goal of this work is therefore to build the structure on which future work can be build upon.

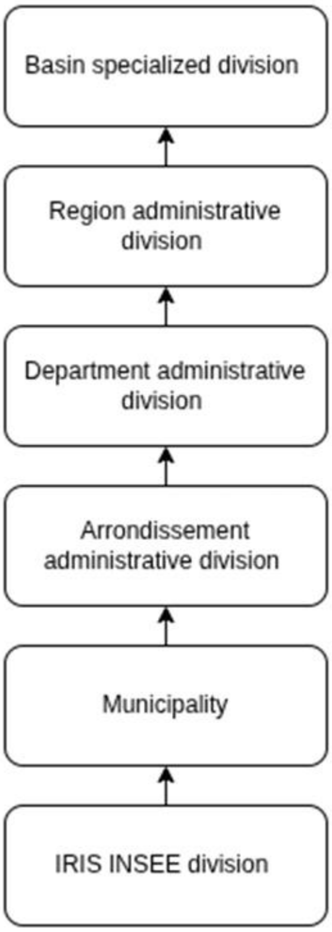


Figure 42: Geospatial architecture used in the upscaled MSM (inspired by the official French administrative territorial divisions)

12.3 Architecture design choices

As mentioned multiple times above, the current simulation model is focused on one tariff which at the moment is defined and enforced at the municipality level. As described in the section “Constraints and challenges for the upscaling operation”, the main aspects of the upscaled model are that it needs to take into account multiple tariffs and with the parameters that go hand-in-hand with it (water providers, demand function, …), as well the spatial dimension problem which can be described as “which municipalities are included in this wider scale ?”.

Since the current simulation model focuses on one tariff and already provides good results for this tariff, it provides a solid starting point for the upscaled model. It would be counterproductive and misguided to start a new model from scratch. Moreover it provides ways to optimize the upscaled model in the future as for example the possibility to retrieve already computed data from a single tariff simulation if the input data for this simulation in specific did not change. This will be explained thoroughly in a later discussion on possible improvements. In summary, the upscaled simulation is built around this existing base simulation model that computes for one tariff and aggregates the results for each tariff to compute the results of a wider entity. This leads to the first two steps of the design process which is : how do we integrate multiple base simulation models in order to gather the results and integrate them.

The first step is trivial in terms of programming but will still be detailed for completeness. The second step is worth spending more time on it.

12.3.1 Base simulation model encapsulation

In order to understand the encapsulation of the base simulation model, it may be worth it to remember how it works strictly in the sense of programming.

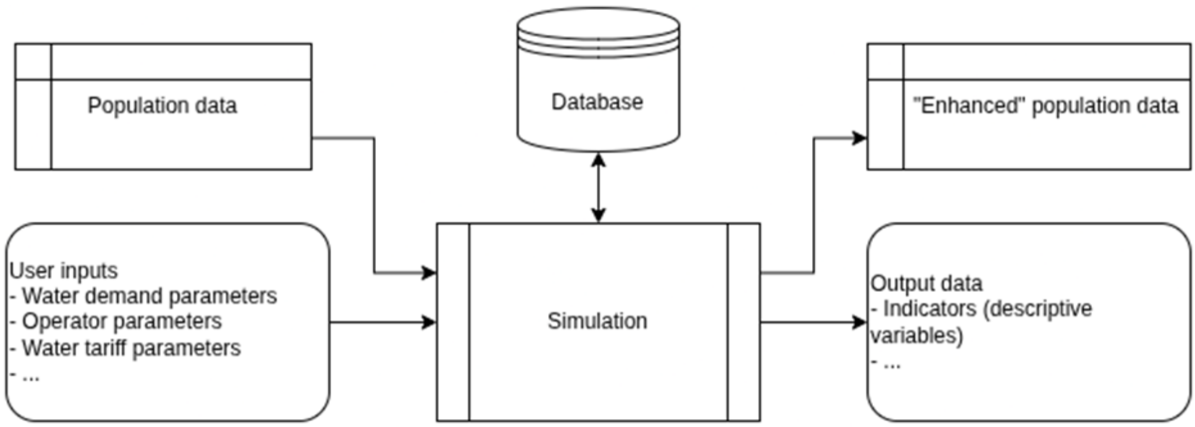


Figure 43: Simplified working principle of the base simulation model

The simulation takes needs to things as input : the population data affected by the tariff, and parameters necessary to derives the indicators (tariff parameters, water operator costs, water demand function, …). For each household and individual in the population data it will first compute extra data needed for the next steps. It then calculates from this data a certain amount of things that can be considered to be descriptive statistics variables as a simplification. Together these variables form the indicators described thoroughly in this present report. These variables and extended population data are stored in the database. In the base simulation model they are also presented in a graphical interface but this part is not important for us at the moment. What is important to know however, is that each of these variables calculations are encapsulated in method calls. In the upscaled simulation model it is unrealistic to call each method for every tariff, hence the need to encapsulate the base simulation model into an object. The basic diagram is given below with the general idea of how the upscaled simulation model would work.

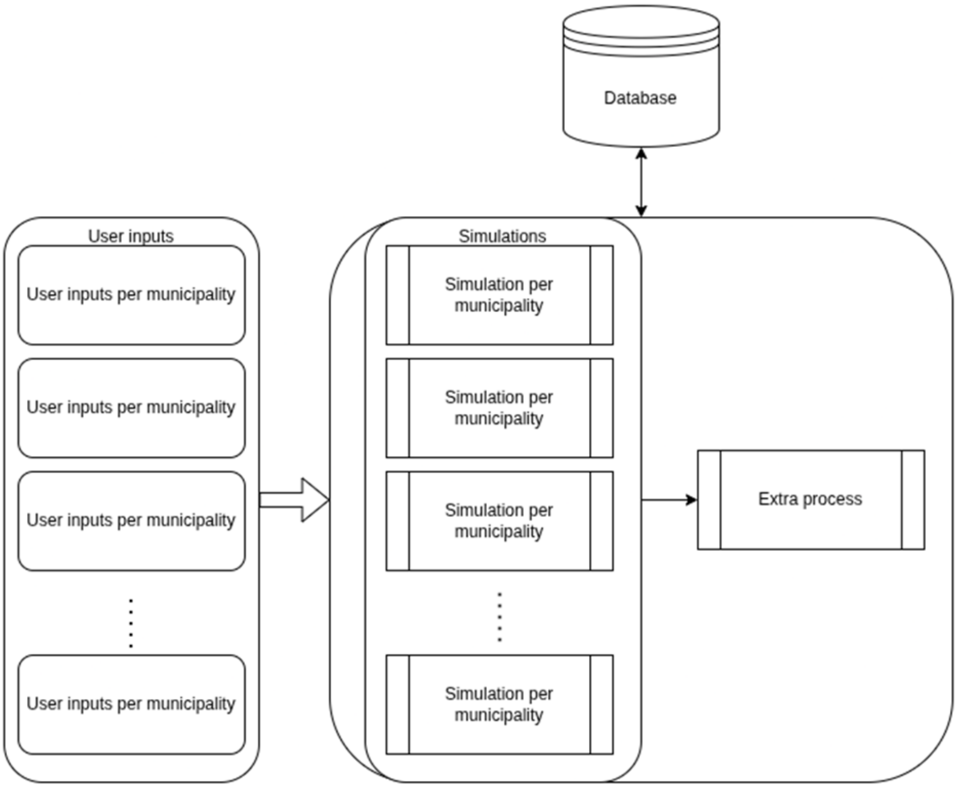


Figure 44: Working principle of the base simulation model encapsulation within the upscaled simulation model

Now that we have hidden the complexity of the base simulation model computation behind an encapsulated model, we can look at what is missing in order to integrate them into the upscaled simulation : how to handle the results in order to achieve the desired results. This unknown is represented by the "extra process" in the simplified diagram.

12.3.2 Population results and aggregators

Once the results are computed by each encapsulated base simulation model, they should be aggregated. This leads to a first problem : not all the results can be aggregated in the same way and more information that just the raw results are needed. For a demonstration purpose, let's take a representative example which is the case of the arithmetic mean aggregation. If we imagine that we are interested in calculating the average income of a population D, which is composed of population A, B and C. The average incomes of each population (A to C) are respectively 2000,2100,2200. The situation is represented in the following table.

Population A income	Population B income	Population C income	Population D income
2000	2100	2200	?

Example table showing the knowledge of averages incomes for three populations A, B and C, together forming the population D

This information alone is not sufficient to calculate the average income of the population D. Indeed in this specific case, we need to compute the weighted average. Thus, we need to know two things : the average incomes we have are **average** descriptive variables, and the population count of each population. We can add this information to the previous table which gives us :

	Population A	Population B	Population C	Population D
Variable name	Income	Income	Income	Income
Variable type	Average	Average	Average	Average
Variable value	2000	2100	2200	?
Population type	Household	Household	Household	Household
Population count N	400	200	400	1000

Example table showing an updated knowledge of an ensemble of populations

Hence, the average income of population D will be weighted average :

$$\begin{aligned} R_{\mathrm{D}} &= \frac{n_{\mathrm{A}}}{n_{\mathrm{D}}} \times R_{\mathrm{A}} + \frac{n_{\mathrm{B}}}{n_{\mathrm{D}}} \times R_{\mathrm{B}} + \frac{n_{\mathrm{C}}}{n_{\mathrm{D}}} \times R_{\mathrm{C}} \\ &= \frac{400}{1000} \times 2000 + \frac{200}{1000} \times 2100 + \frac{400}{1000} \times 2200 \\ &= 2100 \end{aligned}$$

At this stage, it is important to note the “Population type” entry in the table. As the simulation model can focus either on households, individuals or children population inside a municipality population, it is needed to qualify the population we are interested in so that we don’t compare households average income of population A with individuals average income of population B for example, since it would change the population count and skew the results either way.

In the scope of this work, the focus was on developping an architecture that will allow work to continue and to be built upon. To this extend, a choice has been made to integrate two types of results for now : weighted average and variance decomposition. However, it is important to underline that other types of result can be implemented in the same fashion. Since the weighted average has been briefly demonstrated above, the variance decomposition is shown below in order to understand the capabilities of the upscaled simulation model.

$$V(R) = V_{\text{intra}} + V_{\text{inter}} = \sum_h f_h V_h(R) + \sum_h f_h \left(\bar{R}_h - R\right)^2$$

$$f_h = \frac{n_h}{n} \qquad \bar{R}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} R_{ih} \qquad V_h(R) = \frac{1}{n_h} \sum_{i=1}^{n_h} \left(R_{ih} - \bar{R}_h\right)^2$$

Coming back to the weighted average problem and the variance, it leads us to the first design choice : the results are encapsulated in an object called *PopulationResults* in the code. To summarize the previous example in terms of programming language, we went from having enhancing variable name and a variable value to an encapsulated object that contains more information :

- The population of which the result is related
- The population count of which the result is related
- A type in the descriptive statistics sense

Since for a given population it can exist multiple variables that are of the same type such as average variables (e.g. average income, average water consumption, average water invoice, …), they are grouped together in a dictionary. The implemented PopulationResults UML diagram is given below.

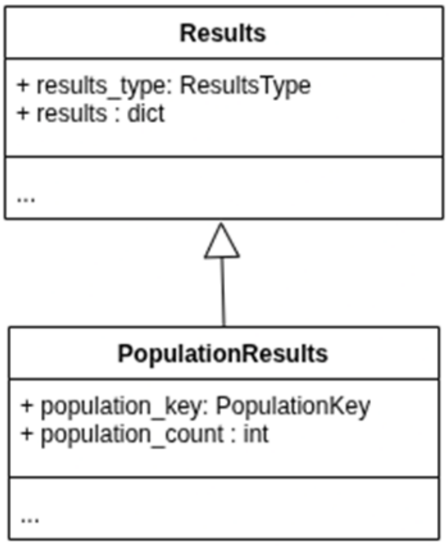


Figure 45 : Simplified UML diagram of the PopulationResults class

It is important to note that PopulationResults is actually defined by a *population_key* instead of a *population_type*. As defined at the moment, a population_key is a tuple that contains two types : PopulationType and SubscriptionType. Indeed, for the needs of the simulation model more flexibility is required in order to define a category of interest to ensure that we compare similar populations. It is also totally possible to extend the population key with other types to better capture a population specificity. A more complete UML diagram is shown below to better understand the PopulationResults concept.

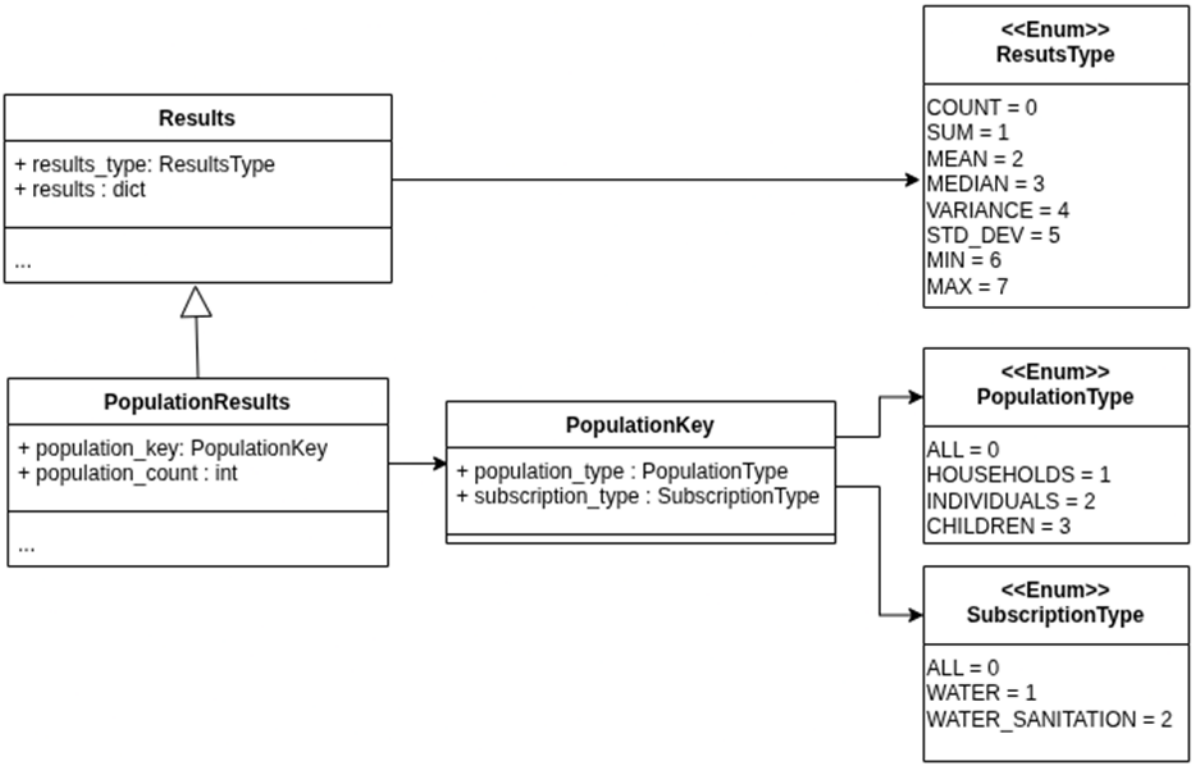


Figure 46: Simplified UML diagram of the PopulationResults class showing the use of PopulationKey and ResultsType

While this is already working great, the amount of resulting objects outputted by each simulation can still be reduced and better organised. Indeed, now each simulation outputs one PopulationResults object per (population_key,result_type) combination. However, most of the time, the results for the same population_key are treated at the same time. It simplifies the code to encapsulate the code in an object called PopulationResultsCollection. The simplified UML diagram is given below :

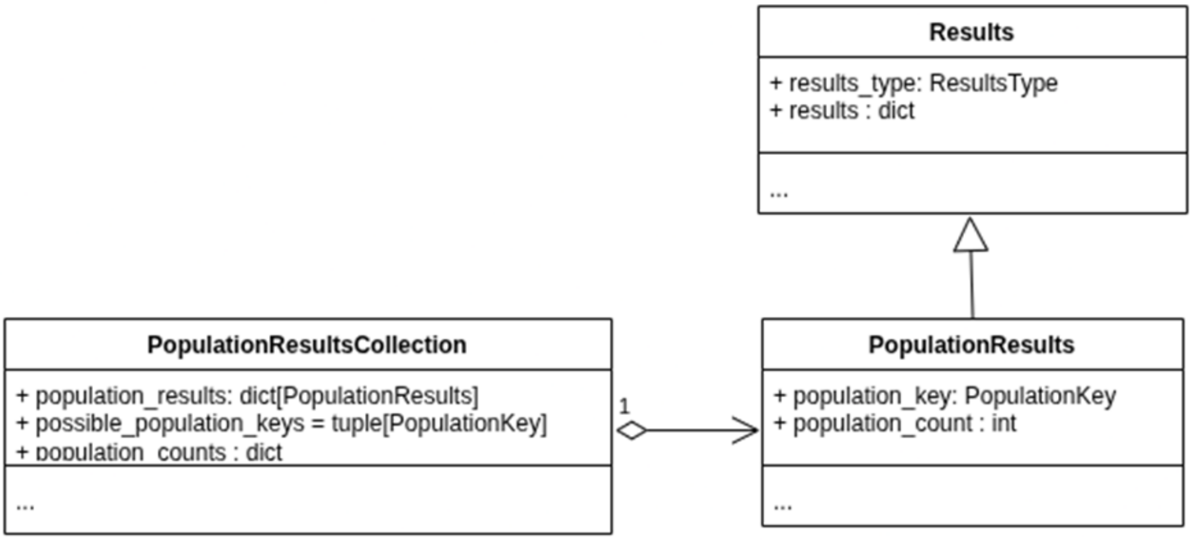


Figure 47: Simplified UML diagram showing the working principle of PopulationResultsCollection

Now that we have defined the output results encapsulation for each simulation, we can introduce the next aspect of the upscaled simulation model : the aggregation.

In order to achieve spatial aggregation, we first need to focus strictly on the aggregation. If we take our example above with the income, the aggregation part is the process of taking PopulationResults A, PopulationResults B and PopulationResults C to output PopulationResults D while calculating the weighted average. Each PopulationResults would be defined by the same PopulationKey (e.g. Households PopulationType with the ALL SubscriptionType) and the same ResultsType (i.e. Mean). This leads to the need of creating an object that would check that the provided PopulationResults match each other and can be aggregated, while using a specific method to do so that depends on the provided ResultsType. This is what the ResultsAggregator object does. A simplified UML diagram is provided below minus some information already provided above.

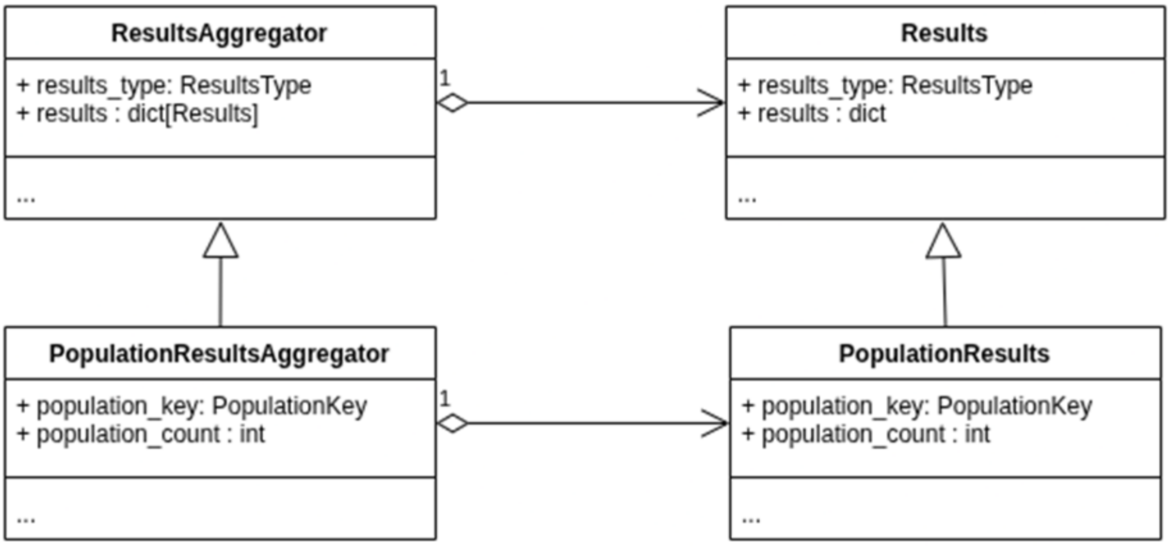


Figure 48: Simplified UML diagram showing the working principle of PopulationResultsAggregator

It is worth noting the population_count attribute in PopulationResultsAggregator. Indeed, it keeps track of the total population that it aggregates in order to provide weighted aggregation methods.

Just in the same flavour that PopulationResultsCollection was used in order to reduce the amount of PopulationResults instance, there is a PopulationResultsCollectionAggregator that is a bit more bloated than all the previous objects. Indeed, it takes a PopulationsResultsCollection as an input and instantiates as many PopulationResultsAggregators as needed depending on all the possible combinations between PopulationKeys available and ResultsType available.

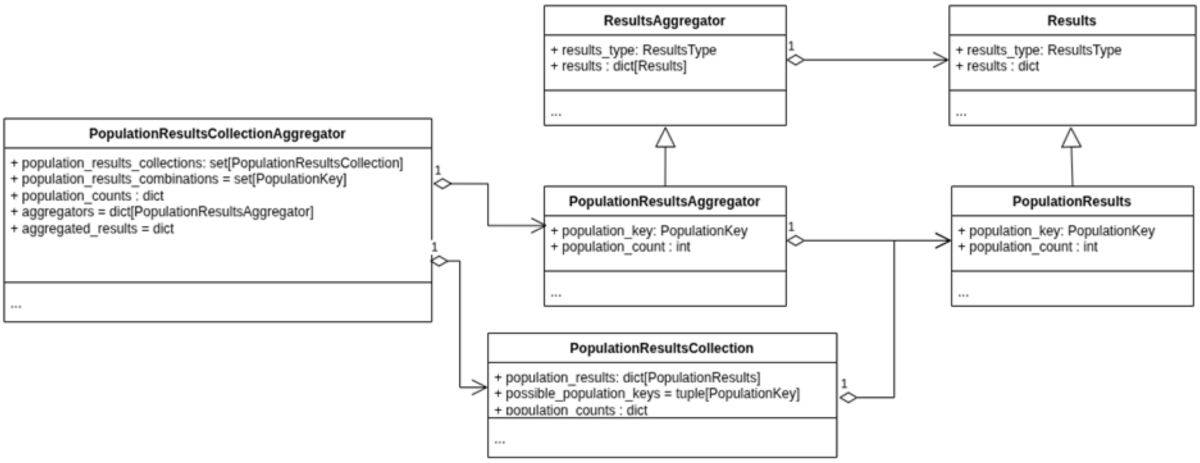


Figure 49: Simplified UML diagram showing the working principle of the PopulationResultsCollectionAggregator

12.3.3 Adding the spatial dimension

Now that we can aggregate PopulationsResults together to get the results of arbitrary grouping, it is time to focus on the spatial aspect of the upscaled simulation model. As a reminder, what we want to aggregate is data from tariff simulations that are defined at the municipality level to another level such as department, arrondissements, etc. It means that at the code level, we need to know what are the municipalities that are included in a target division level. In order to achieve this, the France administrative division from the basin level to the iris level needs to be queryable at the code level. Therefore, the administrative division needs to be stored and represented in a database. It is trivial that SQLAlchemy was used in order to define the relationships between the divisions that are stored in corresponding SQL tables. The SQLAlchemy models that reflect the SQL tables that are shown in the graph Figure 50, on next page.

It is interesting to note here the two tables :

custom_commune_group

commune_group_association.

It was mentioned that a few divisions are not official, and as a matter of fact, customly defined divisions could be created by the user. For now, only custom municipalities grouping can be created but the solution can be used for any type of grouping such as custom_region_groups, etc. The first table custom_commune_group holds the id, the name and the type of the custom municipality group. The type of a group in this case could be inter municipality or micro region (for the case of Réunion Island, it could be North, South, West, East). The table commune_group_association is actually used to hold the references between these custom groups and the involved municipalities.

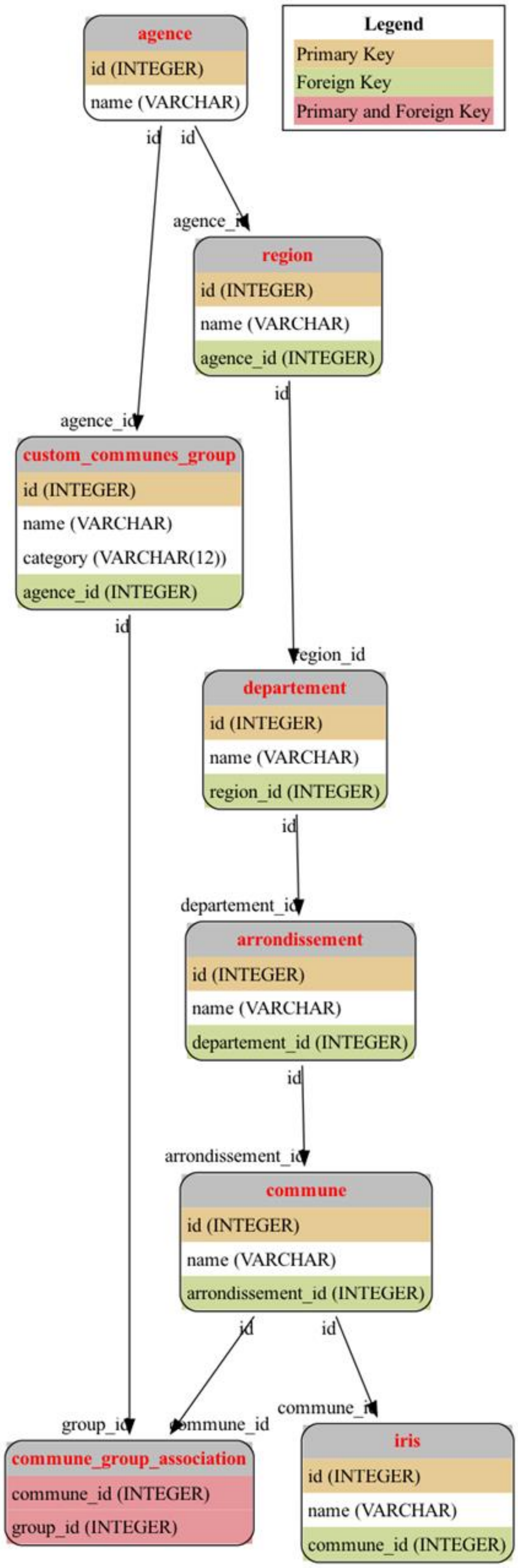


Figure 50: SQL graph showing the structure of the geospatial database

Once again, we have solved a piece of the puzzle and can move on to the next. We have the administrative structure defined in a database, but the next difficult step is to be able to make useful queries out of it. Indeed, the way these are stored is that each child division *"knows"* in which parent division it stands. So for example, it is easy to know from one municipality in which arrondissement it stands, but not directly in which region, and even less to know from the arrondissement id which are the municipalities in it. For this matter, a GeoDBInterface object has been developed. For reference, the "Geo" part stands for Geospatial and allows to make the purpose clear in the code. For example, the GeoDBInterface allows querying for a municipality to which region or department it belongs. It also allows, given a list of *source* division objects and a *target* aggregation level, to output a dictionary with the target aggregation level as keys, and a list of corresponding *source* divisions. Let's give an example to better visualize the aggregation method. Here is a table representing a structure.

<table>
<tr><td colspan="6">Region A</td></tr>
<tr><td colspan="3">Arrondissement B</td><td colspan="3">Arrondissement C</td></tr>
<tr><td>Municipality D</td><td>Municipality E</td><td>Municipality F</td><td>Municipality G</td><td>Municipality H</td><td>Municipality I</td></tr>
</table>

Table 69 : Example of territorial division structure to explain the geospatial aggregation

Without the interface, the only thing that we are able to know is that Municipality G is in Arrondissement C, or that Arrondissement B is in Region A.

If we are interested to group let's say Municipality E, F and H in arrondissements, the GeoDBInterface will return a dictionary : {Arrondissement B : [Municipality E, Municipality F], Arrondissement C : [Municipality H]}.

This will be useful to be coupled with PopulationResultsAggregator to perform the desired calculations. It is useful here to remember that the base simulation model focuses on a tariff and not a municipality. It is not impossible in the future that the simulation could be used outside of France where tariffs are designed at a different level. It can be also imagined that in the future, the responsibility for the tariffs management would shift to higher divisions, such to the arrondissement level. Therefore this explains why the code is generic and allows us to do aggregations at any level.

12.3.4 General upscaled simulation

Now we have all the necessary functionalities to upscale the model, we can describe the general architecture for the upscaled simulation. As mentioned a few times already, the upscaled model runs at its core an encapsulated base simulation model for each tariff. The results are then formatted into PopulationResults objects that are used to sort the results per population key (an arbitrary combination of different population categories) and per result type. The aggregation of the population results is done on a requested administrative division level such as arrondissement or departments. In order to know which tariffs are at the municipal level at the moment, the GeoDBInterface queries the geospatial database in order to provide a dictionary that gathers municipalities together as values associated with the target geospatial level as keys. These groupings of municipalities are used to aggregate the population's results accordingly. It is important to remind, that the geospatial levels at which the data is aggregated is purely arbitrary and chosen by the user. It is not limited to one level and be for example on all the different levels available.

Moreover, it is not impossible that in the future, other processes that are not aggregation could run on the upscaled simulation model just like the base model. This would be highly inefficient but the possibility still exists if some data are absolutely required and can't be aggregated. It is important to know however that it would require to compute again at the population level, defeating the whole point of upscaling.

In order to understand the general structure of the upscaled simulation, a diagram is provided below.

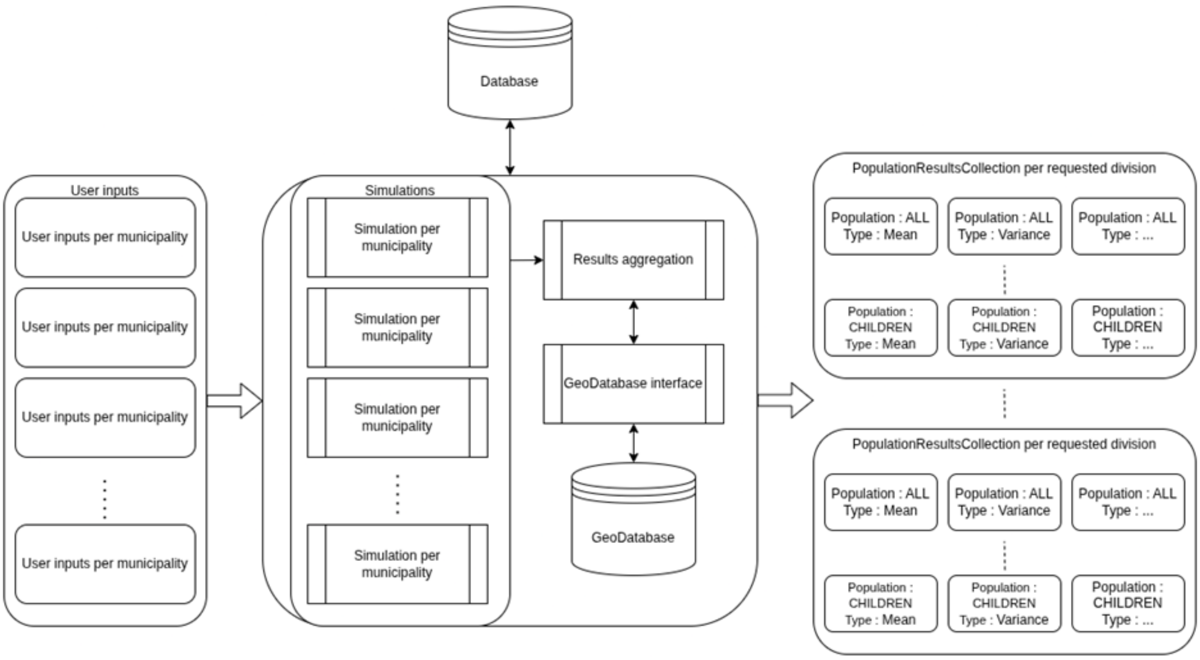


Figure 51: General architecture of the upscaled simulation model.

***Reading tip** : shows how the results coming out of the encapsulated base simulation models are used together with the geospatial knowledge to output data gathered in specific requested territorial divisions*

12.4 Future work

We described what is the current status of the upscaled simulation model. However, there are a few improvements that could be made and some possible solutions will be discussed here.

First, we will focus on the improvements that are relatively easy to implement, referred to as the "low-hanging fruits", which could be put into action in the coming months if more time were available. Next, we will discuss the improvements that were considered but would require more effort and careful planning to implement.

12.4.1 Lowest hanging fruits

The present upscaled simulation shows a prototype of what is feasible, focusing at the moment only on weighted averages and variances decomposition. It is of course possible to increase the amount of results type supported and the code modularity has been designed keeping that in mind. As a brief explanation, the way to add new results types would be to add a category in the ResultsType enumeration, as well as implement the corresponding aggregation method in the Aggregator code. We will here describe a few of these achievable results types that can be added in the very near future.

First of all, the decomposition of the ratios. It is quite obvious that it doesn't present any difficulty. If we consider the percentage of households facing a water abordability issue, to derive the resulting percentage of households of the aggregated population, we need to calculate the weighted average of the percentages. As it has been mentionned, the weighted average has been implemented already, so the work here is minimal.

Next, we consider the Gini index. The work needed to implement this specific result type is a combination between what has been done for the variance. For the within and between components, we can use the same logic as for the variance. However, when it comes to transvariation we can no longer rely on output results given by the encapsulated base simulation model. Indeed, for the transvariation we need to focus on the interaction between two tariff groups (two municipalities in our case), which implies that we need to compare households between two municipalities. It means that we can't rely on the results aggregation and we need to perform similar calculations to what is being done in the base simulation model. It is not impossible and can be perform as a « separate » process from the aggregation. However it is useful to remind that it can dramatically increases the computational power needed and the choice should probably be left to the user appreciation.

As an example, if we consider the basin agency of Loire-Bretagne with around 6500 municipalities, we'd have a matrix with more than 24 400 000 inter municipalities variables for a single Gini index.

Nonetheless, it is important to remind that this is not a difficult process. The upscaled simulation model and the Gini transvariation component calculation has been implemented in the base simulation model.

12.4.2 Ideas for the future

A) Taking advantage of based simulation model encapsulation It was briefly explained that one advantage of base simulation model encapsulation is to optimize the computational power usage. If we take the example of Réunion Island with its 24 municipalities, we could imagine that after a first upscaled simulation we are quite not satisfied with 3 tariffs results. After modifying the tariffs accordingly in these 3 municipalities, we want to run the simulation model again. It would not be efficient at all to recompute the 21 tariffs simulations that did not change. It is entirely possible in the base simulation model to implement before any calculation is done a verification step that checks if the inputs are equal to the inputs of another simulation already computed. If that is the case, the results will be similar and can be provided immediately from the database instead. While we are talking about a few seconds gained per base simulation model, which might not represent a lot for La Réunion, we could be talking about an improvement of minutes per launch of the upscaled simulation model, effectively sparing hours depending on the usage by the operator. We can imagine the same kind of optimisation for the interdependent variables. If you consider municipalities A,B and C where you only modify the input for the municipality C, then you'd get the same results for A, B and A-B. You can gain a lot of time by not recalculating identical results but this checking processes need to be implemented wisely.

B) User interface

1) User inputs As the amount of tariffs and municipalities increases with the amount of parameters and results as well, the readability and usability of the simulation decreases. As a first instance, a spreadsheet is being used to provide the input data to the upscaled simulation. While this is manageable and somewhat readable for the case of La Réunion with its 24 municipalities, it obviously gets more and more difficult with the increase of municipalities. As a reminder and to give an order of magnitude, France has around 35 000 municipalities. It is easy to realize that a spreadsheet with 35 000 rows to enter the tariffs is hard to manage and quickly becomes confusing. Some solutions to make it more manageable can be found such as, for example, providing a spreadsheet with only as many rows as there are modified tariffs compared to a default or saved value. This would reduce the amount of rows and make the spreadsheet more readable and efficient. It obviously is just one example among others.

2) Simulation outputs Unfortunately, the upscaled simulation outputs face the same kind of problems as for the inputs which are related to the sheer amount of data involved in the upscaled simulation. If presenting the indicators and the results for the base simulation model already represented a challenge in order to keep them readable, it is obviously exacerbated by the upscaling operation. While the idea of presenting the results under the form of maps, it still comes with drawbacks. Maps by nature lose information as it transforms the data into colour scales. It also may render outliers less visible if colours are not chosen wisely. Additionally, it would still mean that the upscaled simulation outputs a lot of maps which can be confusing as well if too many are provided.

A proposed solution could be that only a few results interest the user and can be selected beforehand, effectively reducing the amount of data provided.

Moreover, at the moment the upscaled simulation only shows results for the requested levels, but it could also be a possibility that the simulation only shows maps where the tariffs cause problems of inaccessibility with the population, allowing better work and assisting the politics.

3) Takeaways For an uninformed person, it is easy to underestimate the amount of work that the user interface can represent. As it has just been explained, the dramatic increase of input data presents challenges both for the input format as well as the output format.

In summary of the sections about the inputs data and outputs results, the general idea is that there are two main aspect about the work that can be done on the user interface :

- Filtering : The user decides to only focus on some parts of the simulation

- The simulation effectively uses a lot more information that what's being shown, but presents the results in a smart manner to only focus on the problematic parts. The idea is that if a tariff works well, there might be no interest for the user.

The last word is that the user interface is a completely different aspect than what has been done at the moment and necessitates almost a small project on its own. In the end it also depends on the user who is going to use the upscaled simulation model. The core (backend) could be the main focus of future works in order to let each future user work on their own and preferred user interface

C) Geomatic indicators and neighbourhoods matrices The current simulation only aggregates the existing results inside the simulation but does not create new results that inherently come with the geospatial analysis such as heterogeneity and homogeneity of the inaccessibility problem, or the cost of service. For example, it should be the intuition that the more we get away from the water source the higher would the cost of service be. This could be verified quickly by a measure of heterogeneity and an anomaly could be spotted quickly.

Fortunately, these indicators are well known and can be splitted in two groups : indicators that needs to know the neighbourhood relationship matrices between municipalities and indicators that don't.

To this extend, it is important to point out that such neighbourhood matrices are not implemented at the moment. However, it should not be too complicated since it could be implemented as an extra column in the corresponding territorial division SQL table. For example for the case of the municipalities, for each municipality, the id's of the adjacents municipalities can be stored as a list in a column « neighbours_ id ».

A few of these geomatic indicators are given below :

Dimension	Name	Authors
Equality	Segregation index	Duncan and Duncan (1955a,1955b)
	Segregation index adjusted with binary contiguity matrix	Morill (1991)
	Entropy index	Theil (1972); Theil and Finezza (1971)
	Gini index	Duncan and Duncan (1955a)
	Atkinson index	Atkinson (1970)
	…	
Exposition	Isolation index	Bell (1954)
	…	
Concentration	Delta index	Hoover (1941), Duncan et al. (1961)
	Absolute concentration index	Massey and Denton (1988)
	…	
…		

Table 70 : Main geomatic indicators[57]

[57] Extracted from Alivon F [2016]

XIII – CONCLUSION

As shown throughout the document, knowledge of the household water demand functions provided by econometric analysis, coupled with the database used for the econometric estimation (which gives the necessary information on the socio-economic composition of the population living in the area whose the water manager is in charge), provides useful information for measuring the socio-economic performance of the water (and wastewater) pricing policy and identifying some areas for improvement. The information in question relates to the joint distribution within the population of three key factors that are: (1) the volumes of water required to meet basic needs of the households (given the reprocessing of the captive component operated by the user), (2) the sensitivity of demand to changes in tariff parameters and (3) the more or less proper perception of the tariff.

In this respect, it is understood that the issue of the performance of a water and wastewater tariffs needs a detailed answer. Over and above the characteristics of water local demand (that affects the proper design of the tariffs), the impact of a same water pricing policy on affordability, for instance, will be very different depending on whether the poverty rate within the management area is 10% or 40%, even though the household would have the same consumption behaviour, that is would have the same water demand function. In the same vein, the greater or lesser ease of access to available and mobilizable resources, by affecting the cost of the service, also plays a role when it comes to assessing the quality of a pricing policy.

The review of practices shows that water managers (and the community of consultants who support them) do not rely on this tool that constitutes econometrics of water demand when they are called to evaluate their Demand Side Management Policy, with in particular the implementation of the various programme of measures (PGM) and the preparation of the Masterplan for Water Development and Management (SDAGE). This results in a certain weakness in the analyses that are produced today. Thus:

Affordability is measured by calculating the weight of the water bill in household income (CAR for Conventional Affordability Ratio) and a problem is detected when this ratio exceeds a certain threshold, typically 3% according to the recommendations of various international institutes. Calculating this indicator for water managers requires to compute, for each household in their subscriber files, the ratio between the amount of the bill, an information they have, and the customer's income, an information they do not have.

To remedy this difficulty, it is usual to compute the amount of the expense for a reference consumption, typically 120 m³ per year, and determine the level of income below which the CAR exceeds the 3% value. A potential affordability issue is then detected when this threshold income is high, and none is expected to be at stake when it is low[58].

This approach has several limitations.

[58] For example, using the EP tariff for the municipality of Saint Paul (France), calculating the amount of the water bill for a typical consumption of 120 cubic metres per year gives the sum of 46,321 euros per quarter (including VAT), i.e. 15.44 euros per month. Households whose consumption is close to 120 cubic metres with an income of less than $\frac{15.44}{R} = 0.03 \Leftrightarrow R = 514.67$ euros per month are therefore considered to be facing an affordability issue.

- At first, computing the weight of the water bill for a reference consumption of 120 m³ does not necessarily make sense. It is at best representative of an average situation in some contexts and not in others (many regions including Reunion Island have an average consumption greater than 120 m³).

For this reason, the calculation of the income threshold below which the CAR exceeds the 3% threshold value is now often calculated using the average consumption of households as a reference point[59] . However, the main difficulty with this approach is that the households facing an affordability issue are likely to be large poor families whose consumption is large and potentially well above the average.

- In the second place, it is more relevant regarding affordability to work with PAR (Potential Affordability Ratio), defined as the weight in household income of the minimum expenditure on water which is computed by applying the tariff to the volumes of water that are required to meet basic needs.

This indicator can be computed using reference consumption figures, such as those provided by WHO, but one difficulty is that these values are not necessarily representative for the area in which the water manager operates. Indeed, basic needs for water are known to vary through the regions with weather and climatic variables (Rinaudo et al. [2012], Puri & Maas [2020]), household characteristics (Makki et al. [2015], Schleich & Hillenbrand T. [2009]) and quality of the water using equipment (Garcia-Valiñas et al. [2014], Pérez-Urdiales & García-Valiñas [2016], Pérez-Urdiales et al. [2016]). Moreover, the lack of information on household income still prevents the water manager from computing the value of PAR for each household in its customer file and thus taking the actual measure of the problem within the population the water company is in charge.

Alternately, having an econometric model of (local) household water demand, coupled with the database used to derive it[60], makes it possible to estimate the distribution of basic needs within the customer file and, in so doing, estimate the distribution of PAR within the population of the (local) households, with a number of additional analyses that can be conducted. In particular, it is possible to:

- compute impact indicators (such as the average of excessive charges)

and, based on the determinants of basic consumption (provided for a part by the econometric model),

- identify the socio-economic profiles of households that are facing unaffordability,
- identify the consumption factors (including the characteristics of the water tariff) that makes these households falling into unaffordability.

---

[59] For instance, the average domestic water consumption in the commune of Saint Paul is 240 cubic metres per year, i.e. 60 cubic metres per quarter, with a water bill of 63.362 euros per quarter (including VAT) and a household income threshold of $\frac{63.362}{3R} = 0.03 \Leftrightarrow R = 704.02$ per month. The interpretation is therefore the same as above, except that it is for households that are close to the average consumption value, which is also particularly high (double the national average).

[60] And which can always be obtained by surveying subscribers. It should also be emphasised that the use of an econometric model can be carried out in strict compliance with regulations on the protection of personal data, with the anonymisation of household data that feeds the Population module.

In addition,

- the possibility is given to simulate the effects on all these items of a change in tariff parameters (with all the usual precautions), including the characteristics of the first consumption blocks when an IBT applies.

All these elements (some of which are implemented by the tool) are clearly useful to inform public decision making in this area.

The second item is in line with the EU-WFD and focuses on measurement of the incentive effect of the pricing system, that is its ability to fix households on water-efficient and saving uses. To gauge the latter, it is therefore expected that the various indicators that are used will rely on measures of overconsumption. In fact, given the lack of information on the latter, the practice consists of:

- computing an indicator called the "average price", defined as the amount of the water bill divided by the level of water consumption,
- identifying the tariff threshold beyond which this average price increases with consumption (if any),
- comparing the various IBTs on the basis of this tariff threshold beyond which the tariff is regarded as progressive

(see, in particular, the example of EP tariffs for the communes of Saint Denis and Saint Paul, given page 40 and the Figure 9 & Figure 10 page 41). The lower the threshold value, the greater the incentive.

The underlying idea behind the calculation of this indicator is to consider that subscribers react to the average price, including the amount of the subscription[61], and that they will perceive an IBT as being a degressive tariff, encouraging them to consume more in the first consumption blocks, where the average price is locally decreasing.

While this model (which is implicitly mobilised when this indicator is used) is sustainable from a theoretical point of view, the main difficulty is that it is not linked to the measurement of overconsumption. Alternately, with an econometric model of local household water demand, there is a room for:

- estimating the distribution of overconsumption within the household customer file,
- assessing the impact of changes in tariff parameters on this distribution of overconsumption,

---

[61] The question of whether or not to include the subscription fee in the calculation of the price perceived is clearly an open question that needs to be answered empirically. In all cases, it should be emphasised that the ironing behaviour of households, whether it concerns the graph of the tariff function or the graph of the unit price scale, is not systematic. For example, when it comes to pricing methods for Internet and mobile phone services, which often consist of a single fixed charge (or, more precisely, presented as such), it is not uncommon for households to state that their consumption costs them nothing. While this statement is strictly speaking not correct - in all cases there is payment of an access fee - households do identify the price they pay for the consumption of an additional unit, which is effectively zero in this case. This response is therefore entirely consistent with a Nordin-style specification for which there is an initial consumption block characterised by a marginal price equal to 0 (these flat-rate systems are in fact often progressive tariffs, particularly in the area of Internet subscriptions, with a relatively high threshold for the first block to cover basic usage).

- identifying in this overconsumption a part linked to poor perception of the tariff (allocative inefficiency) and computing the induced costs of poor management that are borne by the households.

Here again, all of these elements that are implemented by the tool are useful for assessing, as far as possible, the performance of the tariff in this incentive dimension. It should also be kept in mind that IBT scheme, because it presents a certain degree of complexity, is often poorly perceived by the households and may lead to significant overconsumption.

In fact, almost all the analyses carried out by water managers to assess the socio-economic performance of the water pricing policy they implement are limited to the two practices just described (at least, for the most French cities).

It should also be noted that, due to a lack of information on the properties of water demand (the values of the price elasticities in particular), the stakeholders involved in water pricing setting have little insight into the financial consequences of revising water tariffs in a context where, moreover, water and sanitation services are often the subject of a public service delegation agreement.

In this context, the poor socio-economic performance of current IBTs enlightened by academic literature (presented in the general introduction to this document) makes the think that water managers and public decision-makers lack visibility on the socio-economic consequences of the pricing policies they implement, and even on the features of the decision problem they face to and that may go against their intuition (such as, for instance, the fact that the correlation between water consumption and household income is perceived to be strong while it is empirically weak).

The tool which aims to enable to:

(i) make a clear diagnosis of the gaps, on each of the 5 items related to EU-WFD,

(ii) identify through simulation exercises some potential points of improvement,

(iii) highlight the nature and quantify the trade-offs between the various objectives the water pricing policy has to meet,

focusses specifically on these lacks. It is hoped that the further developments of the tool (that needs to be improved on numerous aspects) can effectively be implemented in the coming years to enable it to be disseminated on a wide scale, firstly to stakeholders and, potentially, also to citizens.

XIV – ANNEXES

Annex 1: Laws of Demand – Nordin reading

Taking the equation (5.8) given on page 60, the conditional demand of a household located in Block 2 is given by :

$$\begin{aligned}\ln q_i &= \ln q_{0i} - 0.31 \ln \pi_2 + 0.25 \ln\left(R_i - F + D_2\right) \\ &= \ln q_{0i} - 0.31 \ln \pi_2 + 0.25 \ln\left(R_i - F + \left(\pi_2 - \pi_1\right) k_1\right)\end{aligned} \quad (14.1)$$

Looking at this formulation, it appears that the price of Block 1, π_1, and the threshold of Block 1, k_1, have an effect on the conditional demand of Block 2 through an income effect and the change in Nordin's D, $D_2 = (\pi_2 - \pi_1)k_1$, that decreases with π_1 and increases with k_1 :

$$-\left.\frac{\partial \ln q_i}{\partial \pi_1}\right|_{q_i \in I_2} = \frac{0.25}{R_i - F + D_2} \times \frac{\partial D_2}{\partial \pi_1} = -\frac{0.25}{R_i - F + D_2} \times k_1 < 0 \quad (14.2)$$

$$-\left.\frac{\partial \ln q_i}{\partial k_1}\right|_{q_i \in I_2} = \frac{0.25}{R_i - F + D_2} \times \frac{\partial D_2}{\partial k_1} = -\frac{0.25}{R_i - F + D_2} \times \left(\pi_2 - \pi_1\right) > 0 \quad (14.3)$$

Concerning now the price variation for Block 2:

$$-\left.\frac{\partial \ln q_i}{\partial \pi_2}\right|_{q_i \in I_2} = -\frac{0.31}{\pi_2} + \frac{0.25}{R_i - F + D_2} \times \frac{\partial D_2}{\partial \pi_2} = -\frac{0.31}{\pi_2} + \frac{0.25}{R_i - F + D_2} \times k_1 \quad (14.4)$$

that is the price of the consumption block in which the household is assumed to be located, the impact combines:

(1) a direct effect that pulls consumption down,

(2) an indirect effect, linked to the increase in the Nordin's D, $D_2 = (\pi_2 - \pi_1)k_1$, that drives consumption up (trough an income effect).

With a few calculations, it can be shown that the first effect outweighs numerically the second one, with a total effect that is ultimately negative[62] . Moreover, since the conditional demand function (14.1) does not depend on the price of block 3, the price of block 4 …, nor the tariff thresholds k_2, k_3 …, it appears that variations in these parameters have no impact (locally).

Focussing now on the case of a consumer located in block 3 for whom:

$$\begin{aligned}\ln q_i &= \ln q_{0i} - 0.31 \ln \pi_3 + 0.25 \ln\left(R_i - F + D_3\right) \\ &= \ln q_{0i} - 0.31 \ln \pi_3 + 0.25 \ln\left(R_i - F + \left(\pi_2 - \pi_1\right) k_1 + \left(\pi_3 - \pi_2\right) k_2\right)\end{aligned} \quad (14.5)$$

[62] In demand theory, this income effect linked to a price variation generally emerges as soon as the agent's dotations are impacted by this variation. For this reason, the negative impact of the marginal price π_j on the conditional demand of Block *j* can be considered as "usual".

it appears that:

- the price of Block 2 will now have an impact on water demand through an income effect:

$$-\left.\frac{\partial \ln q_i}{\partial \pi_2}\right|_{q_i \in I_3} = \frac{0.25}{R_i - F + D_3} \times \frac{\partial D_3}{\partial \pi_2} = -\frac{0.25}{R_i - F + D_2} \times (k_2 - k_1) < 0 \qquad (14.6)$$

- what will henceforth play the role of price is the price of Block 3:

$$-\left.\frac{\partial \ln q_i}{\partial \pi_3}\right|_{q_i \in I_3} = -\frac{0.31}{\pi_3} + \frac{0.25}{R_i - F + D_3} \times \frac{\partial D_3}{\partial \pi_3} = -\frac{0.31}{\pi_3} + \frac{0.25}{R_i - F + D_3} \times k_2 < 0 \qquad (14.7)$$

(marginal price) with an effect that proves to be ultimately negative;

- the impact of the Block 2 threshold, given its effet on the value of Nordin's D (of Block 3), is no longer equal to 0 with:

$$-\left.\frac{\partial \ln q_i}{\partial k_2}\right|_{q_i \in I_3} = \frac{0.25}{R_i - F + D_3} \times \frac{\partial D_3}{\partial k_2} = -\frac{0.25}{R_i - F + D_3} \times (\pi_3 - \pi_2) > 0 \qquad (14.8)$$

The effects of π_1 and k_1 are, moreover, similar to those given by (14.2) and (14.3) once the denominator evaluated at $R - F + D_3$.

Note The calculation of the impact:

$$-\left.\frac{\partial \ln q_i}{\partial \pi_h}\right|_{q_i \in I_j} = -\left.\left(\frac{1}{q_i}\frac{\partial q_i}{\partial \pi_h}\right)\right|_{q_i \in I_j} \qquad (14.9)$$

for *h* varying from 1 to *p* and *i* varying from 1 to *n*, take precise numerical values for each household in the Population file, given the features of the tariff that is evaluated/tested by the user. This information can be used to assess the impact of a small change in price π_h by applying the following formulae:

$$\frac{\Delta q_i}{q_i} \approx \left.\frac{\partial \ln q_i}{\partial \pi_h}\right|_{q_i \in I_j} \times \Delta \pi_h \qquad (14.10)$$

that gives the percentage variation in consumption of household *i* following an increase in the price of Block *h* of $\Delta \pi_h$ euros (to a first approximation). Similarly, the impact of a small variation in the block threshold k_h can be assessed by making use of a similar relation that writtes:

$$\frac{\Delta q_i}{q_i} \approx \left.\frac{\partial \ln q_i}{\partial k_h}\right|_{q_i \in I_j} \times \Delta k_h \qquad (14.11)$$

The values obtained can be interpreted as giving (to a first approximation) the variation in household *i*'s consumption, in %, following an increase in the threshold k_h of Δk_h cubic metres.

Annex 2: Usual statistical indicators - methodological elements

With regard to the interpretation of basic statistical indicators, it should be remembered that the arithmetic mean of a statistical series gives the value of the variable of interest when all individuals are identical / have the same value of the characteristic (with heterogeneity then equal to 0). In concrete terms, the figure of €1.39 for the IBT reported in Table 5 indicates that, if all households faced the same affordability deficit, each household would overpay by €1.39 per quarter on its water bill (which may or may not include a payment for collective sanitation). Compared with the TBSE, for which this average value is €17.47 per quarter (literally, if all households faced the same affordability problem, each household would pay the sum of €17.47 too much, per quarter, with the implementation of the TBSE), the implementation of the IBT therefore results in a substantial reduction in unaffordability (which is divided by 12.6 compared with its reference value). Naturally, this assessment of the performance of the IBT needs to be refined on a number of points.

The first is that, because it is calculated on the household population as a whole, this figure includes in its scope of calculation the population of household subscribers not facing an affordability issue, i.e. those for whom the affordability deficit is equal to 0. It is then natural to complete the value of the average by calculating an average for the sole population of households facing an affordability issue, equal in this case to 17.69 euros for the IBT and 53.33 euros for the TBSE. Literally:

- If all households facing an affordability issue with the IBT that is assessed/tested by the user faced an issue of the same magnitude, all these households (that do face an issue) would overpay €17.69 per quarter on their water bills (the latter may or may not include a payment for collective sanitation),

and, compared with the TBSE for which this average / identically distributed value is €53.33 per quarter, the impact on the actual affordability deficit, while still significant, is now only 3.0 (compared with 12.6 for the apparent affordability deficit).

The second factor is that the comparison of effective averages for IBT and TBSE is based on populations of different sizes, unlike the average for apparent affordability (for which the field is identical). In this respect, it is important to be able to assess the tariff's ability to lift households out of unaffordability by comparing its heacount ratio with its TBSE reference value. Literally, the comparison of the IBT and TBSE figures shows that the IBT tested/evaluated here by the user takes $32.8 - 7.9 = 24.9\%$ households out of unaffordability, while reducing the amount of excessive charges by $53.33 - 17.69 = 35.64$ euros per quarter for households that remain unaffordable.

Remark Households that are unaffordable under the IBT may not be unaffordable under the TBSE. In other words, the implementation of a IBT may result in households becoming unaffordable that were not initially unaffordable with the TBSE. This could be the case for households with high basic consumption levels who would be faced with an IBT for which the subsidised consumption blocks are small and the tax (margin) rates on high consumption levels are high. For more advanced processing of the data, the net number of exits from unaffordability should therefore be broken down into its two gross components: the number of exits from unaffordability minus the number of entries into unaffordability. In practice, as these cases are rare, particularly in view of the fact that subsidies on the right of access are often quite high, the tool does not display any particular information on this item, and the user (who always has the option of exporting the data) is advised to supplement these figures with additional analyses relating to the socio-economic composition of households facing an affordability issue (what the tool is doing by providing in the third stage of the analysis, the percentage of poor households within the group of households facing an affordability issue).

The third element is that these two factors both play a role in the (in this case significant) fall in the apparent affordability deficit, which is itself equal to the product of the Heacount ratio of households multiplied by the effective affordability deficit (see equations (7.5) et (7.6)). One can quantify the contributions of these two factors to reducing unaffordability by approximating the percentage fall in the affordability deficit by the sum of the (here negative) growth rate of the Household Headcount ratio (extensive dimension) and the (here negative) growth rate of the actual affordability deficit (intensive dimension). Given its approximate nature (with an approximation that is all the more satisfactory the smaller the relative variations), this (simple) manipulation is left to the care (and appreciation) of the user.

Next, having established these first trends, it is clear that not all households face the same affordability problem. In particular, there is heterogeneity (which is totally masked by the average figure) and this needs to be measured. To do this, one possible approach is to measure the distance between the statistical series of (apparent / effective) unaffordability available to us, for example :

Next, having established these initial findings, it is clear that not all households face the same affordability issue. In particular, there is heterogeneity (which is masked by the average figure) that needs to be measured. To do this, one possible approach is to measure the distance between the statistical series of (apparent / effective) unaffordability, for instance:

$$\mathbf{e} = \left(e_1^{\text{PAR\_IBT}}, e_2^{\text{PAR\_IBT}}, \ldots, e_n^{\text{PAR\_IBT}}\right) \tag{14.12}$$

(PAR IBT unaffordability series) and a theoretical one in which all individuals would be identical, and present an affordability deficit equal to the average unaffordability:

$$\overline{\mathbf{e}}_{\text{PAR\_IBT}} = \left(\overline{e}_{\text{PAR\_IBT}}, \overline{e}_{\text{PAR\_IBT}}, \ldots, \overline{e}_{\text{PAR\_IBT}}\right) \tag{14.13}$$

with :

$$\overline{e}_{\text{PAR\_IBT}} = \frac{1}{n}\sum_{i=1}^{n} e_i^{\text{PAR\_IBT}} = \frac{1}{n}\sum_{i=1}^{n} \max\left[T_{\text{IBT}}\left(\underline{q}_i\right) - 3\%R_i, 0\right] \tag{14.14}$$

(and a user value of 3% for the threshold above which the household is considered to be facing an affordability issue for the PAR criterion). For these purposes, a natural response is to calculate the average of deviations from the arithmetic mean (MAPE):

$$\text{MAPE}\left(\mathbf{e}_{\text{PAR\_IBT}}\right) = \frac{1}{n}\sum_{i=1}^{n}\left|e_i^{\text{PAR\_IBT}} - \overline{e}_{\text{PAR\_IBT}}\right| \tag{14.15}$$

that is the value of the distance (in euros) separating each household from the average unaffordability if all households were at the same distance from the average unaffordability[63] . A

[63] Based on a population of 100 households, half would be below average, the other half above.

second indicator, spontaneously less natural but systematically used, consists of calculating the root mean square of the deviations from the arithmetic mean or **standard deviation**:

$$\sigma\left(\mathbf{e}_{\mathrm{PAR\_IBT}}\right)=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(e_i^{\mathrm{PAR\_IBT}}-\overline{e}_{\mathrm{PAR\_IBT}}\right)^2} \quad (14.16)$$

also equal to the square root of the variance:

$$V\left(\mathbf{e}_{\mathrm{PAR\_IBT}}\right)=\frac{1}{n}\sum_{i=1}^{n}\left(e_i^{\mathrm{PAR\_IBT}}-\overline{e}_{\mathrm{PAR\_IBT}}\right)^2 \quad (14.17)$$

(itself defined as the average of the squares of the deviations from the arithmetic mean). These two indicators, MAPE and standard deviation, refer to two different notions of distance, namely the Taxi distance :

$$d_{\mathrm{t}}\left(\mathbf{e}_{\mathrm{PAR\_IBT}},\overline{\mathbf{e}}_{\mathrm{PAR\_IBT}}\right)=\left|e_1^{\mathrm{PAR\_IBT}}-\overline{e}_{\mathrm{PAR\_IBT}}\right|+\left|e_2^{\mathrm{PAR\_IBT}}-\overline{e}_{\mathrm{PAR\_IBT}}\right|+\ldots+\left|e_n^{\mathrm{PAR\_IBT}}-\overline{e}_{\mathrm{PAR\_IBT}}\right| \quad (14.18)$$

and the usual (Euclidean) distance :

$$d\left(\mathbf{e}_{\mathrm{PAR\_IBT}},\overline{\mathbf{e}}_{\mathrm{PAR\_IBT}}\right)=\sqrt{\left(e_1^{\mathrm{PAR\_IBT}}-\overline{e}_{\mathrm{PAR\_IBT}}\right)^2+\left(e_2^{\mathrm{PAR\_IBT}}-\overline{e}_{\mathrm{PAR\_IBT}}\right)^2+\ldots+\left(e_i^{\mathrm{PAR\_IBT}}-\overline{e}_{\mathrm{PAR\_IBT}}\right)^2}$$

(14.19)

to measure the proximity between 2 points (in this case of dimension n) which are the series of unaffordability actually available:

$$\mathbf{e}=\left(e_1^{\mathrm{PAR\_IBT}},e_2^{\mathrm{PAR\_IBT}},\ldots,e_n^{\mathrm{PAR\_IBT}}\right) \quad (14.20)$$

and the one that would appear if everyone was identical:

$$\overline{\mathbf{e}}_{\mathrm{PAR\_IBT}}=\left(\overline{e}_{\mathrm{PAR\_IBT}},\overline{e}_{\mathrm{PAR\_IBT}},\ldots,\overline{e}_{\mathrm{PAR\_IBT}}\right) \text{ avec } \overline{e}_{\mathrm{PAR\_IBT}}=\frac{1}{n}\sum_{i=1}^{n}e_i^{\mathrm{PAR\_IBT}} \quad (14.21)$$

i.e. if the household heterogeneity of unaffordability were zero. As apparent, we have indeed:

$$d_t\left(\mathbf{e}_{\mathrm{PAR\_IBT}},\overline{\mathbf{e}}_{\mathrm{PAR\_IBT}}\right)=n\times\mathrm{MAPE}\left(\mathbf{e}_{\mathrm{PAR\_IBT}}\right)$$

$$d\left(\mathbf{e}_{\mathrm{PAR\_IBT}},\overline{\mathbf{e}}_{\mathrm{PAR\_IBT}}\right)=\sqrt{n}\times\sigma\left(\mathbf{e}_{\mathrm{PAR\_IBT}}\right)$$

Literally, MAPE and standard deviation measure, equal up to a constant, the Taxi/Euclidean distance between the actual series and the series in which everyone would be identical. As such, these indicators are effectively, in the first sense of the term, a measure of household heterogeneity concerning the affordability of the EP / EPA tariff system.

As a general rule, the standard deviation indicator is preferred, notably because it is linked to the variance, which has the advantage of being decomposable into the sum of intra-group heterogeneity and inter-group heterogeneity (unlike MAPE, which is not). As this variance decomposition is implemented at a later stage in the analysis, the tool displays also the value of variance to make the link with subsequent processing (but, in terms of interpretation, the value of the standard deviation, whose unit of measurement here is in euros, is to be preferred (unit of measurement if variance is here in euros squared)). Data from Table 5 show that, compared with the TBSE, the implementation of the IBT leads to reductions in individual heterogeneity in terms of unaffordability both in the general population (with a MAPE that falls from €24.49 to

€2.57 per quarter and a standard deviation which falls from €31.83 to €6.21 per quarter) as well as for the only population affected by unaffordability (the MAPE then falls from €30.19 to €11.25 per quarter and the standard deviation from €34.17 to €13.22 per quarter).

Finally, and to conclude with these methodological aspects (which govern the selection and interpretation of these basic descriptive statistics indicators relating to this first level of analysis), it should be noted that variations in the mean play a role in variations in the standard deviation (or variance) in a context where, strictly speaking, the comparison of standard deviations (or variances) is fully satisfactory only when the two statistical series have the same mean (otherwise, the difference in means also plays a role in the difference in variances). It is then possible to take account of this effect (on heterogeneity) linked to the variation in the mean by calculating and comparing the coefficients of variation of the variable of interest, defined as the ratio of the standard deviation (here, of affordability) to the mean (here, of affordability). This comparison is based on a specific operation ("data resizing") which is presented in the box below. In this case, it shows that, in relative terms, i.e. in relation to the average for the situation under consideration, the (relative) gaps have increased with the switch from TBSE to IBT. One can use this information to state that, based on a constant mean, the affordability gaps would be greater with the IBT than with the TBSE.

Box - the coefficient of variation Consider 2 series of unaffordability:

$$\mathbf{x} = (x_1, x_2, \dots, x_n)$$

$$\mathbf{y} = (y_1, y_2, \dots, y_n)$$

with, for example, the first that refers to IBT unaffordability while the second refers to TBSE unaffordability with $\bar{x} < \bar{y}$ (see above). Next, to compare individual heterogeneity (variance, standard deviation) with a constant mean, the data needs to be rescaled so that the two series have the same mean. To do so, one possible operation is to multiply all the affordability deficits x_i of the series $\mathbf{x} = (x_1, x_2, \dots, x_n)$ by the same coefficient $k = \bar{y} / \bar{x} > 1$, to obtain a new series:

$$\mathbf{x}' = (x'_1, x'_2, \dots, x'_n) = (k \times x_1, k \times x_2, \dots, k \times x_n) = (\frac{\bar{y}}{\bar{x}} \times x_1, \frac{\bar{y}}{\bar{x}} \times x_2, \dots, \frac{\bar{y}}{\bar{x}} \times x_n)$$

(multiplicative transformation) which has the same mean as the series $\mathbf{y} = (y_1, y_2, \dots, y_n)$:

$$\bar{x}' = \frac{1}{n}\sum_{i=1}^{n} x'_i = \frac{1}{n}\sum_{i=1}^{n} k \times x_i = k \times \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = k \times \bar{x} = \frac{\bar{y}}{\bar{x}} \times \bar{x} = \bar{y}$$

It is also to note that, with this operation on the data, the relative structure of unaffordability within the series $x' = (x'_1, x'_2, \dots, x'_n)$, by verifying:

$$\frac{x'_i}{x'_j} = \frac{k \times x_i}{k \times x_j} = \frac{x_i}{x_j} \text{ for all } i = 1, \dots, n \text{ and } j = 1, \dots, n$$

is not affected by this operation, i.e. if the affordability deficit of household *i* is initially twice as large as the affordability deficit of household *j*, the affordability deficit of household *i* is still twice as large as that of household *j* after the data has been rescaled. Once this rescaling operation carried out, it is now possible to make a constant mean comparison of the variances of the affordability deficits of the series $\mathbf{x}' = k \cdot \mathbf{x}$ et y, $V(\mathbf{x}')$ and $V(\mathbf{y})$. Given the properties of the variance, it is known (i) that the heterogeneity of the affordability deficits in the rescaled series $x' = k \cdot x$ has been multiplied by k^2 with here:

$$V(\mathbf{x}') = V(k \cdot \mathbf{x}) = k^2 \times V(\mathbf{x}) = \left(\frac{\bar{y}}{\bar{x}}\right)^2 \times V(\mathbf{x}) > V(\mathbf{x})$$

(because $k > 1 \Leftrightarrow \bar{y} > \bar{x}$) with (ii) a variance ratio equal to:

$$\frac{V(\mathbf{x}')}{V(\mathbf{y})} = \frac{k^2 \times V(\mathbf{x})}{V(\mathbf{y})} = k^2 \times \frac{V(\mathbf{x})}{V(\mathbf{y})} = \left(\frac{\bar{y}}{\bar{x}}\right)^2 \times \frac{\sigma_{\mathbf{x}}^2}{\sigma_{\mathbf{y}}^2} = \left(\frac{\bar{y}}{\bar{x}}\right)^2 \times \left(\frac{\sigma_{\mathbf{x}}}{\sigma_{\mathbf{y}}}\right)^2 = \left(\frac{\sigma_{\mathbf{x}}}{\bar{x}}\right)^2 \times \frac{1}{\left(\sigma_y / \bar{y}\right)^2} = \frac{cv_{\mathbf{x}}^{\;2}}{cv_{\mathbf{y}}^2}$$

therefore equal to :

$$\frac{V(\mathbf{x}')}{V(\mathbf{y})} = \left(\frac{cv_{\mathbf{x}}}{cv_{\mathbf{y}}}\right)^2$$

Under these conditions, the heterogeneity within the series x can be described as relatively stronger (respectively relatively weaker) if its coefficient of variation :

$$cv_{\mathbf{x}} = \frac{\sigma_{\mathbf{x}}}{\bar{x}}$$

is higher (respectively lower) than that of the y series:

$$cv_{\mathbf{y}} = \frac{\sigma_{\mathbf{y}}}{\bar{y}}$$

In addition, we have :

$$\frac{V(\mathbf{x})}{V(\mathbf{y})} = \frac{V(\mathbf{x})}{V(\mathbf{x}')} \times \frac{V(\mathbf{x}')}{V(\mathbf{y})} = \frac{V(\mathbf{x})}{k^2 \times V(\mathbf{x})} \times \left(\frac{cv_{\mathbf{x}}}{cv_{\mathbf{y}}}\right)^2 = \frac{1}{k^2} \times \left(\frac{cv_{\mathbf{x}}}{cv_{\mathbf{y}}}\right)^2 = \left(\frac{\bar{x}}{\bar{y}}\right)^2 \times \left(\frac{cv_{\mathbf{x}}}{cv_{\mathbf{y}}}\right)^2$$

$$\Leftrightarrow V(\mathbf{x}) = \left(\frac{\bar{x}}{\bar{y}}\right)^2 \times \left(\frac{cv_{\mathbf{x}}}{cv_{\mathbf{y}}}\right)^2 \times V(\mathbf{y})$$

$$\Leftrightarrow \sigma_x = \frac{\bar{x}}{\bar{y}} \times \frac{cv_{\mathbf{x}}}{cv_{\mathbf{y}}} \times \sigma_{\mathbf{y}}$$

This last relationship establishes (i) that the direct comparison of standard deviations is indeed biased by the fact that the two series do not have the same mean and, once corrected for this bias, (2) that the ratio of the coefficients of variation makes it possible to infer (and quantify) the increase in the heterogeneity of the unaffordability between the two series $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$. ■

Annex 3: Group Decomposition of Gini index

A3.1 Gini Index – A reminder

A) Notations The following notations are used:

(1) $P = \{1, 2, \ldots, n\}$ the population of statistical units (individuals, households, etc.) indicated by the letter *i*,

(2) x_i the value taken by the variable of interest, in this case the amount of unaffordability (also known as the affordability deficit) for household *i*,

(3) $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ the series of affordability deficits in the population *P*, ordered from smallest to largest value with $0 \le x_1 \le x_2 \le \ldots \le x_n$, also known as the Parade de Pen (of the affordability deficit).

(4) $X = x_1 + x_2 + \ldots + x_n$ the total affordability deficit,

(5) $F_i = i / n$ the position of agent *i* (or normalized rank or fractional rank), *i* varying from 1 to n, in Pen's parade $\mathbf{x} = (x_1, x_2, \ldots, x_n)$,

(6) $\mu = \overline{x}$ the average affordability deficit of the population P :

$$\mu = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{1}{n} \times \sum_{i=1}^{n} x_i$$

therefore, the value of the affordability deficit if all households faced the same affordability deficit, taking into account the value of total unaffordability X (equal sharing of a "pie" of size X);

(7) α_i the share of total unaffordability $X = x_1 + x_2 + \ldots + x_n$ supported by household *i* :

$$\alpha_i = \frac{x_i}{x_1 + x_2 + \ldots + x_n} = \frac{x_i}{\sum_{i=1}^{n} x_i} \qquad (14.22)$$

for *i* varying from 1 to n,

(8) A_i the cumulative sum of the shares α_1, ..., α_i :

$$A_i = \alpha_1 + \alpha_2 + \ldots + \alpha_i = \frac{x_1 + x_2 + \ldots + x_i}{x_1 + x_2 + \ldots + x_i + \ldots + x_n} = \frac{x_1 + x_2 + \ldots + x_i}{n\mu} \qquad (14.23)$$

for *i* varying from 1 to n,

(9) $A_i = L(F_i)$ the Lorenz curve associated to Pen's parade $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ (in fact, its linear extension),

(10) i the Gini index (approximately twice the area of concentration) of unaffordability for the general population P. See Figure 52, page 254.

B) Calculating the Gini index One possible method (referred as the "trapezoidal method") consists in calculating :

$$i = 1 - \sum_{i=1}^{n} \frac{A_i + A_{i-1}}{n} \qquad (14.24)$$

The secund proceeds with the following formula:

$$i = \frac{\Delta}{2\overline{x}} = \frac{1}{2\overline{x}} \times \frac{1}{n^2} \sum_{i=1}^{n} \sum_{i'=1}^{n} |x_i - x_{i'}| \qquad (14.25)$$

with Δ the average of inter-individual gap (here, of affordability deficits).

Numerical example For Pen's parade of unaffordability $\mathbf{x} = (x_1, x_2, x_3) = (100, 200, 300)$, with $x_1 = 100$ the affordability deficit of household 1, $x_2 = 200$ the affordability deficit of household 2 and $x_3 = 300$ the affordability deficit of household 3. Table 3.1 is drawn upon first. The Lorenz curve (shown in Figure 52, page 254) is constructed from points $(0,0)$, $(F_1, A_1) = (\frac{1}{3}, \frac{1}{6})$, $(F_2, A_2) = (\frac{2}{3}, \frac{2}{6})$ and $(F_3, A_3) = (1,1)$, and the calculation of the Gini index (using the trapezoid method) gives $i = \frac{2}{9} = 22.2\%$. This value is also equal to that obtained by applying formula (14.25) to the matrix of inter-individual differences in affordability given in Table 3.2 with :

$$\Delta = \frac{1}{n^2} \times \sum_{i=1}^{n} \sum_{i'=1}^{n} |x_i - x_{i'}| = \frac{0+100+200+100+0+100+200+100+0}{3^2} = \frac{800}{9}$$

and $i = \frac{\Delta}{2\overline{x}} = \frac{1}{2 \times 200} \times \frac{800}{9} = \frac{2}{9}$.

i	x_i	f_i	F_i	α_i	A_i	A_{i-1}	$f_i \times (A_i + A_{i-1})$
1	100	1/3	1/3	1/6	1/6	0	$\frac{1}{3} \times (0 + \frac{1}{6}) = \frac{1}{18}$
2	200	1/3	2/3	2/6	3/6	1/6	$\frac{1}{3} \times (\frac{3}{6} + \frac{1}{6}) = \frac{2}{9}$
3	300	1/3	1	½	1	3/6	$\frac{1}{3} \times (1 + \frac{3}{6}) = \frac{1}{2}$
Σ	600	1					$\frac{7}{9}$
	$\overline{x} = \frac{600}{3} = 200$						$i = 1 - \frac{7}{9} = \frac{2}{9}$

Table 3.1 : Calculation of the Gini index using the trapezoid method - example

$\lvert x_i - x_{i'} \rvert$	1 (100)	2 (200)	3 (300)
1 (100)	$\lvert 100-100 \rvert = 0$	$\lvert 100-200 \rvert = 100$	$\lvert 100-300 \rvert = 200$
2 (200)	$\lvert 200-100 \rvert = 100$	$\lvert 200-200 \rvert = 0$	$\lvert 200-300 \rvert = 100$
3 (300)	$\lvert 300-100 \rvert = 200$	$\lvert 300-200 \rvert = 100$	$\lvert 300-300 \rvert = 0$

Table 3.2 : Calculating the Gini index from the matrix of inter-individual gaps (affordability deficits) – example

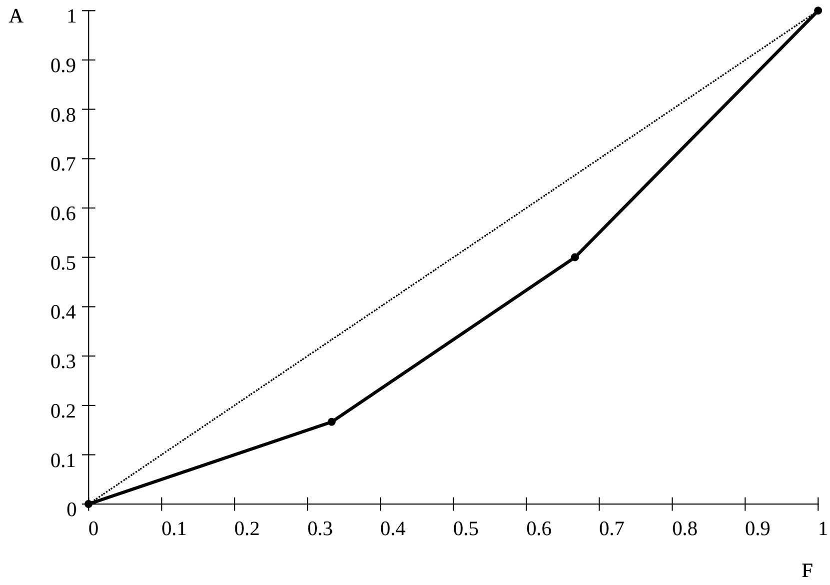


Figure 52

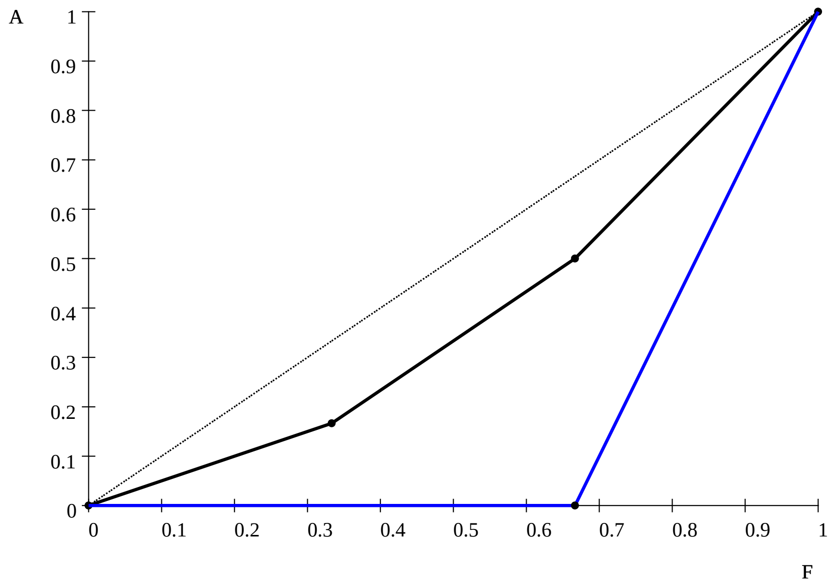


Figure 53

C) Properties In the discrete case, (i) the maximum value of the Gini index (which occurs when "one agent has everything, the others have nothing") is given by $i = 1 - \frac{1}{n}$ with $i \to 1$ as $n \to +\infty$, and (ii) the minimum value (which occurs in the egalitarian case) is given by 0 (with a Lorenz curve which coincides with the 45° line / the equality line).

Numerical Illustration – continued Starting from Pen's parade $\mathbf{x} = (x_1, x_2, x_3) = (100, 200, 300)$, one transfers the affordability deficits of households 1 and 2 to household 3 to obtain the distribution $\mathbf{x}' = (x_1', x_2', x_3') = (0, 0, 600)$. Voir alors (i) Figure 53, page 254, with the Lorenz curve that changes (blue curve), (ii) Table 3.1 that becomes:

i	x_i	f_i	F_i	α_i	A_i	A_{i-1}	$f_i \times (A_i + A_{i-1})$
1	100	1/3	1/3	0	0	0	$\frac{1}{3} \times (0+0) = 0$
2	200	1/3	2/3	0	0	0	$\frac{1}{3} \times (0+0) = 0$
3	300	1/3	1	1	1	0	$\frac{1}{3} \times (1+0) = \frac{1}{3}$
Σ	600	1					$\frac{1}{3}$
	$\overline{x} = \frac{600}{3} = 200$						$i = 1 - \frac{1}{3} = \frac{2}{3}$

and (iii) the new value of the Gini index, which is equal to $1 - \frac{1}{n} = 1 - \frac{1}{3} = \frac{2}{3}$. In the other polar case, which is that of equality, obtained by transferring the 100 of unaffordability from household 3 to household 1, we have $\mathbf{x}'' = (x_1'', x_2'', x_3'') = (200, 200, 200)$ and Table 3.1 that becomes:

i	x_i	f_i	F_i	α_i	A_i	A_{i-1}	$f_i \times (A_i + A_{i-1})$
1	200	1/3	1/3	1/3	1/3	0	$\frac{1}{3} \times (\frac{1}{3} + 0) = \frac{1}{9}$
2	200	1/3	2/3	1/3	2/3	1/3	$\frac{1}{3} \times (\frac{2}{3} + \frac{1}{3}) = \frac{1}{3}$
3	200	1/3	1	1/3	1	2/3	$\frac{1}{3} \times (1 + \frac{2}{3}) = \frac{5}{9}$
Σ	600	1					$\frac{1}{9} + \frac{1}{3} + \frac{5}{9} = 1$
	$\overline{x} = \frac{600}{3} = 200$						$i = 1 - 1 = 0$

and the Gini index that is equal to 0. As apparent, the Lorenz curve (in fact, its extension by linearity) merges with the 45° line in this case (with the area of concentration that is equal to 0). For the sake of completeness, it should be noted that, in the polar case where inequalities are maximal, the table of inter-individual gap is of the form :

$\lvert x_i - x_{i'} \rvert$	1 (0)	2 (0)	3 (600)
1 (0)	$\lvert 0-0 \rvert = 0$	$\lvert 0-0 \rvert = 0$	$\lvert 0-600 \rvert = 600$
2 (0)	$\lvert 0-0 \rvert = 0$	$\lvert 0-0 \rvert = 0$	$\lvert 0-600 \rvert = 600$
3 (600)	$\lvert 600-0 \rvert = 600$	$\lvert 600-0 \rvert = 600$	$\lvert 600-600 \rvert = 0$

with therefore:

$$i_2 = \frac{\Delta_2}{2\overline{x}_2} = \frac{1}{2\times 200}\times\frac{0+0+600+0+0+600+600+600+0}{3^2} = \frac{1}{2\times 200}\times\frac{2400}{9} = \frac{6}{9} = \frac{2}{3}$$

At the opposite, if total unaffordability (of 600) is distributed equally among the 3 agents, the table of inter-individual gap is of the form:

$\lvert x_i - x_{i'}\rvert$	1 (200)	2 (200)	3 (200)
1 (200)	$\lvert 200-200\rvert = 0$	$\lvert 200-200\rvert = 0$	$\lvert 200-200\rvert = 0$
2 (200)	$\lvert 200-200\rvert = 0$	$\lvert 200-200\rvert = 0$	$\lvert 200-200\rvert = 0$
3 (200)	$\lvert 200-200\rvert = 0$	$\lvert 200-200\rvert = 0$	$\lvert 200-200\rvert = 0$

with $\Delta_3 = 0$ and $i_3 = \Delta_3 / 9 = 0$. ■

A3.2 Decomposition of Lambert & Aronson [1993]

A) Additional assumptions and notations The population P (of n households) is divided into k subgroups P_1, P_2, …, P_k indicated by the letter j with:

(11) n_j the size of group *j* ;

(12) f_j the weight of sub-population *j* in population P :

$$f_j = \frac{n_j}{n} \tag{14.26}$$

(13) $\mathbf{x}_j = (x_{j1}, x_{j2}, \ldots, x_{jn})$ the Pen's parade of unaffordability within group *j*,

(14) $\overline{x}_j = \mu_j$ the average unaffordability deficit of group *j* :

$$\mu_j = \overline{x}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} x_{ji} = \frac{X_j}{n_j} \tag{14.27}$$

with x_{ji} the affordability deficit of household *i* member of the group *j* and:

$$X_j = \sum_{i=1}^{n_j} x_{ij} = n \times \mu_j \tag{14.28}$$

the total affordability deficit of group *j*,

(15) α_j the share of the affordability deficit supported by group *j* :

$$\alpha_j = \frac{X_j}{X} = \frac{n_j \times \mu_j}{n \times \mu} = f_j \times \frac{\mu_j}{\mu} \tag{14.29}$$

(16) i_j the Gini index of group *j*.

B) Identifying the within component It is to identify the contribution of within-group inequality to total inequality as measured by the Gini index:

$$i = \frac{\Delta}{2\overline{x}} = \frac{1}{2\overline{x}} \times \frac{1}{n^2} \sum_{i=1}^{n} \sum_{i'=1}^{n} \left| x_i - x_{i'} \right|$$

The starting point is the table of inter-individual affordability gaps on the next page, in which individuals are ranked (i) in ascending order of the average affordability of the group to which they belong, i.e. in ascending order of average affordability deficits $\mu_1 \le \mu_2 \le \ldots \le \mu_k$, next (ii) in ascending order of the affordability deficit within their group they belong (lexico-graphic Pen's parade). On this basis, the affordability deficit gaps $\left| x_i - x_{i'} \right|$ for individuals i and i' belonging to the same group j, $j = 1, \ldots, k$ are identified with therefore:

$$i_w = \frac{1}{2\overline{x}} \times \frac{1}{n^2} \sum_{j=1}^{k} \sum_{i=1}^{n_j} \sum_{i'=1}^{n_j} \left| x_{ji} - x_{ji'} \right| \tag{14.30}$$

("w" pour within). Next, we have:

$$\begin{aligned} i_w &= \frac{1}{2\overline{x}} \times \frac{1}{n^2} \sum_{j=1}^{k} \sum_{i=1}^{n_j} \sum_{i'=1}^{n_j} \left(\frac{n_j^2}{n_j^2} \times \frac{\mu_j}{\mu_j} \times \left| x_{ji} - x_{ji'} \right| \right) = \sum_{j=1}^{k} \sum_{i=1}^{n_j} \sum_{i'=1}^{n_j} \left(\frac{n_j^2}{n_j^2} \times \frac{\mu_j}{\mu_j} \times \frac{\left| x_{ji} - x_{ji'} \right|}{2n^2 \times \overline{x}} \right) \\ &= \sum_{j=1}^{k} \sum_{i=1}^{n_j} \sum_{i'=1}^{n_j} \left(\frac{n_j^2 \times \mu_j}{n^2 \times \overline{x}} \times \frac{\left| x_{ji} - x_{ji'} \right|}{2n_j^2 \times \mu_j} \right) = \sum_{j=1}^{k} \sum_{i=1}^{n_j} \sum_{i'=1}^{n_j} \left(\frac{n_j}{n} \times \frac{n_j \times \mu_j}{n \times \overline{x}} \times \frac{\left| x_{ji} - x_{ji'} \right|}{2n_j^2 \times \mu_j} \right) \\ &= \sum_{j=1}^{k} \sum_{i=1}^{n_j} \sum_{i'=1}^{n_j} \left(f_j \times \frac{X_j}{X} \times \frac{\left| x_{ji} - x_{ji'} \right|}{2n_j^2 \times \mu_j} \right) = \sum_{j=1}^{k} \sum_{i=1}^{n_j} \sum_{i'=1}^{n_j} \left(f_j \alpha_j \times \frac{\left| x_{ji} - x_{ji'} \right|}{2n_j^2 \times \mu_j} \right) \\ &= \sum_{j=1}^{k} f_j \alpha_j \times \sum_{i=1}^{n_j} \sum_{i'=1}^{n_j} \left(\frac{\left| x_{ji} - x_{ji'} \right|}{2n_j^2 \times \mu_j} \right) = \sum_{j=1}^{k} f_j \alpha_j \times \left(\frac{1}{2n_j^2 \times \mu_j} \sum_{i=1}^{n_j} \sum_{i'=1}^{n_j} \left| x_{ji} - x_{ji'} \right| \right) \\ &= \sum_{j=1}^{k} f_j \alpha_j \times i_j \end{aligned} \tag{14.31}$$

with:

$$i_j = \frac{\Delta_j}{2\overline{x}_j} = \frac{1}{2\overline{x}_j} \times \frac{1}{n_j^2} \sum_{i=1}^{n} \sum_{i'=1}^{n} \left| x_{ji} - x_{ji'} \right| \tag{14.32}$$

the Gini index for group *j* (calculated from the matrix of inter-individual differences for group *j*). In fine, the intra component of the Gini index is measured by :

	$1:1$	$1:2$	$\cdots$	$1:n_1$	$2:1$	$2:2$	$\cdots$	$2:n_2$	$\cdots$
$1:1$	$\lvert x_{11}-x_{11}\rvert=0$	$\lvert x_{11}-x_{12}\rvert$		$\lvert x_{11}-x_{1n_1}\rvert$	$\lvert x_{11}-x_{21}\rvert$	$\lvert x_{11}-x_{22}\rvert$		$\lvert x_{11}-x_{2n_2}\rvert$	$\cdots$
$1:2$	$\lvert x_{12}-x_{11}\rvert$	$\lvert x_{12}-x_{12}\rvert=0$		$\lvert x_{12}-x_{1n_1}\rvert$				$\lvert x_{12}-x_{2n_2}\rvert$	$\cdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$	
$1:n_1$	$\lvert x_{1n_1}-x_{11}\rvert$	$\lvert x_{1n_1}-x_{12}\rvert$		$\lvert x_{1n_1}-x_{1n_1}\rvert=0$	$\lvert x_{1n_1}-x_{21}\rvert$	$\lvert x_{1n_1}-x_{22}\rvert$		$\lvert x_{1n_1}-x_{2n_2}\rvert$	$\cdots$
$2:1$	$\lvert x_{21}-x_{11}\rvert$	$\lvert x_{21}-x_{12}\rvert$		$\lvert x_{21}-x_{1n_1}\rvert$	$\lvert x_{21}-x_{21}\rvert=0$	$\lvert x_{21}-x_{22}\rvert$		$\lvert x_{21}-x_{2n_2}\rvert$	
$2:2$	$\lvert x_{22}-x_{11}\rvert$	$\lvert x_{22}-x_{12}\rvert$		$\lvert x_{22}-x_{1n_1}\rvert$	$\lvert x_{22}-x_{21}\rvert$	$\lvert x_{22}-x_{22}\rvert=0$		$\lvert x_{22}-x_{2n_2}\rvert$	$\cdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\cdots$
$2:n_2$	$\lvert x_{2n_2}-x_{11}\rvert$	$\lvert x_{2n_2}-x_{12}\rvert$		$\lvert x_{2n_2}-x_{1n_1}\rvert$	$\lvert x_{2n_2}-x_{12}\rvert$	$\lvert x_{2n_2}-x_{22}\rvert$		$\lvert x_{2n_2}-x_{2n_2}\rvert=0$	
$\vdots$									

Table 71 : Matrix of inter-individual gaps in affordability deficit with group ranking, next individual ranking ("lexico-graphic Pen's Parade")

$$i_w = \sum_{j=1}^{k} w_j \times i_j \tag{14.33}$$

with $w_j = f_j \alpha_j$, *j* varying from 1 to k, the weight to be assigned to the Gini index of group *j*, i_j, in calculating the within component i_w of Gini index i for population P as a whole. It is to note that the within component of Gini index is written as a weighted sum (and not a weighted average) of the various Gini index of sub-groups P_1, P_2, …, P_k.

B) Identifying the between component The difference $i - i_w$ measures, by construction, the proportion of inequalities that is not due to intra-group inequalities. This non-intra component proves to be the sum of two other components, including an inter-group component i_b ("b" pour between) linked to the gaps in average unaffordability deficit between the groups μ_1, μ_2, …, μ_k. This between component is simple to calculate. It corresponds to the Gini index calculated on the "Between" Pen's parade:

$$\begin{aligned}\mathbf{B} &= (x_{11}, x_{12}, \ldots, x_{1n_1}, x_{21}, x_{22}, \ldots, x_{2n_2}, \ldots, x_{k1}, x_{k2}, \ldots, x_{k,n_k}) \\ &= (\mu_1, \mu_1, \ldots, \mu_1, \mu_2, \mu_2, \mu_2 \ldots, \mu_2, \ldots, \mu_k, \mu_k, \mu_k, \ldots, \mu_k)\end{aligned}$$

with $\mu_1 \le \mu_2 \le \ldots \le \mu_k$ the average affordability deficits of the sub-groups, ranked from the lowest one to the highest one.

C) Identifying the transvariation component The second part of the non-intra component $i - i_w$ is the Transvariation component i_t, linked to the overlap of Pen's parade of unaffordability of sub-groups:

$$\mathbf{x}_1 = (x_{11}, x_{12}, \ldots, x_{1n})$$

$$\mathbf{x}_2 = (x_{21}, x_{22}, \ldots, x_{2n})$$

…

$$\mathbf{x}_k = (x_{k1}, x_{k2}, \ldots, x_{kn})$$

Since the sub-group decomposition of the Gini index is exact with:

$$i = i_w + i_b + i_t \tag{14.34}$$

this Transvariation component can be computed by balance (this result is due to Lambert & Aronson [1993]).

D) Visualising the decomposition Following the "Introducing inequality in stages" methodology of Lambert [2001], one can visualise the decomposition of the Gini index (which measures the transition from the equality line to the original Lorenz curve) using the following procedure:

Step 1 Starting from a situation of equality where all households face the same affordability issue :

$$\begin{aligned}\boldsymbol{\mu} &= (x_{11}, x_{12}, \ldots, x_{1n_1}, x_{21}, x_{22}, \ldots, x_{2n_2}, \ldots, x_{k1}, x_{k2}, \ldots, x_{k,n_k}) \\ &= (\mu, \mu, \ldots, \mu, \mu, \mu, \ldots, \mu, \ldots, \mu, \mu, \ldots, \mu)\end{aligned} \tag{14.35}$$

(with a Lorenz curve that merges with the 45° line), inter-group inequalities are introduced (i) by assigning to each individual *i* the mean of its reference group μ_j, (ii) by ordering the *n* individuals according to the sub-group average values $\mu_1 \le \mu_2 \le \ldots \le \mu_k$ to get the "Between" Pen's Parade:

$$\begin{aligned}\mathbf{B} &= (x_{11}, x_{12}, \ldots, x_{1n_1}, x_{21}, x_{22}, \ldots, x_{2n_2}, \ldots, x_{k1}, x_{k2}, \ldots, x_{k,n_k}) \\ &= (\mu_1, \mu_1, \ldots, \mu_1, \mu_2, \mu_2, \ldots, \mu_2, \ldots, \mu_k, \mu_k, \ldots, \mu_k)\end{aligned} \tag{14.36}$$

and (iii) to construct the related Lorenz curve, $A = L_b(F)$. The Gini index associated with this B Lorenz curve :

$$i_b = \int_0^1 \left(F - L_b(F)\right) dF \tag{14.37}$$

corresponds to the Between component of the original Gini index.

Step 2 Starting from the "Between" Pens Parade, individual heterogeneity is introduced by ranking individuals within their reference group according to the value of their personal affordability deficit to obtain the lexico-graphic Pen's Parade:

$$\mathbf{Lex} = (x_{11}, x_{12}, \ldots, x_{1n_1}, x_{21}, x_{22}, \ldots, x_{2n_2}, \ldots, x_{k1}, x_{k2}, \ldots, x_{k,n_k}) \tag{14.38}$$

and, next, the graph of the unaffordability concentration curve, $A = L_{lex}(F)$, that is related to this specific lexico-graphical ranking of household. The difference between (i) the area of concentration associated with the graph $A = L_{lex}(F)$ and (ii) the area of concentration associated with the graph of $A = L_b(F)$, that is :

$$\int_0^1 \left(F - L_{lex}(F)\right) dF - \int_0^1 \left(F - L_b(F)\right) dF = \int_0^1 \left(L_b(F) - L_{lex}(F)\right) dF \tag{14.39}$$

measures the within component, denoted i_w, of the Gini index i.

Etape 3 Strating from the lexico-graphic Pen's Parade and the related concentration curve $A = L_{lex}(F)$, some households in groups with the lowest average affordability deficits may have an affordability deficit greater than those of some households in groups with the highest average unaffordability deficit. Reranking households according to the value of their personal unaffordability deficit leads to the original Pen parade $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and the related original Lorenz curve $A = L(F)$. The difference in areas:

$$\int_0^1 \left(F - L(F)\right) dF - \int_0^1 \left(F - L_{lex}(F)\right) dF = \int_0^1 \left(L_{lex}(F) - L(F)\right) dF \tag{14.40}$$

accounts for the impact on the Gini index of this re-ranking of households to return to the original Pen's parade / Lorenz curve and, as such, accounts for the Transvariation component of the Gini index i_t, linked to overlaps in the conditional distributions of unaffordability deficit. See Figure 54, on the next page (the variable of interest is the CAR affordability deficit in the general population (0 included), with $i = 0.827$, $i_b = 0.370$, $i_w = 0.356$, $i_w = 0.356$ and $i_t = 0.101$).

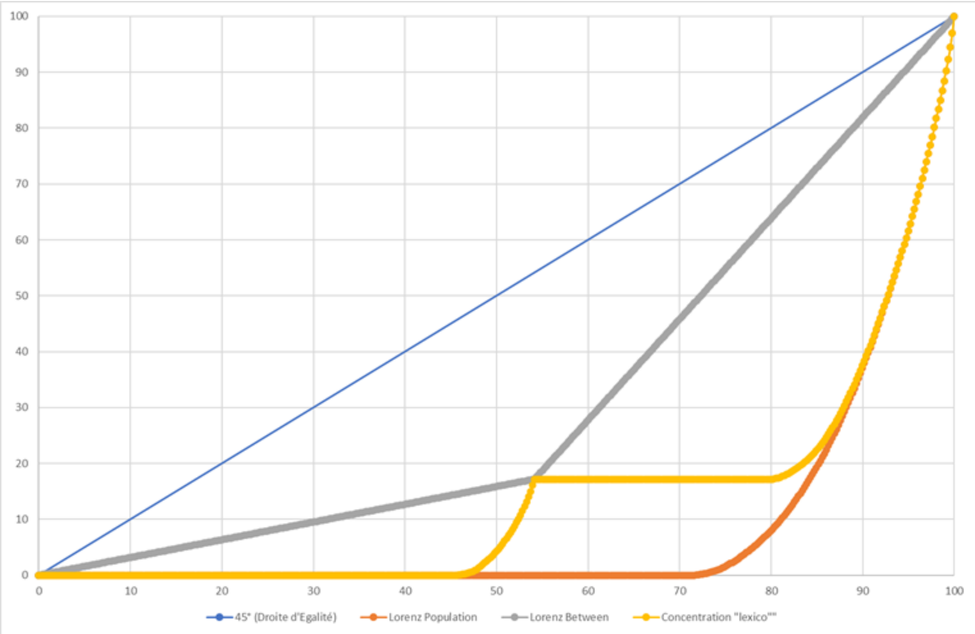


Figure 54 : : breakdown of the Gini index of Apparent CAR Unaffordability (0 included

E) Numerical example 2 groups of of 3 individuals with affordability deficits set to 100, 200 and 300 for the first subrgoup, and to 200, 400 and 600 for the second subgroup, are considered. Households are indiced by 1, 2 and 3 for group 1, and by 1', 2' and 3' for group 2. The calculation of the Gini index for the population as a whole is performed as follows:

i	x_i	f_i	F_i	α_i	A_i	A_{i-1}	$f_i \times (A_i + A_{i-1})$
1	100	1/6	1/6	1/18	1/18	0	$\frac{1}{6} \times (\frac{1}{18} + 0)$
2	200	1/6	2/6	2/18	3/18	1/18	$\frac{1}{6} \times (\frac{3}{18} + \frac{1}{18})$
1’	200	1/6	3/6	2/18	5/18	3/18	$\frac{1}{6} \times (\frac{5}{18} + \frac{3}{18})$
3	300	1/6	4/6	3/18	8/18	5/18	$\frac{1}{6} \times (\frac{8}{18} + \frac{5}{18})$
2’	400	1/6	5/6	4/18	12/18	8/18	$\frac{1}{6} \times (\frac{12}{18} + \frac{8}{18})$
3’	600	1/6	1	6/18	1	12/18	$\frac{1}{6} \times (1 + \frac{12}{18})$
Σ	1800	1		1			$\frac{19}{27}$
	$\overline{x} = \frac{1800}{6}$	$= 300$					$i = \frac{8}{27} = 0.296$

Table 3.3

One will note that individuals are ranked according to the value of their affordability deficit, from the smallest one to the largest one, with individual 3 in group 1 having an affordability deficit greater than the one of individual 1' in group 2, in a framework where the affordability deficit of group 2 is, on average, greater than the average affordability deficit of group 1. In addition;

(i) **Calculation of the Gini index for group 1** The distribution of unaffordability for group 1, $\mathbf{x}_1 = (100, 200, 300)$, is identical to the one considered above. We have therefore $i_1 = 0.222$.

(ii) **Calculation of the Gini index for group 2** As shown in Table 3.4, we have $i_2 = 0.222$ with, therefore, $i_2 = i_1$, what is an expected result because (i) the distribution of group 2 satisfies $\mathbf{x}_2 = (200, 400, 600) = 2 \cdot \mathbf{x}_1 = 2 \cdot (100, 200, 300)$, and (ii) it proves that the value of Gini index does not change when the values on which it is calculated change proportionnaly[64].

(2) **Calculation of the Between component** The affordability distributions for groups 1 and 2 are given by $\mathbf{x}_1 = (100, 200, 300)$ and $\mathbf{x}_2 = (200, 400, 600)$, with $\mu_1 = 200$ et $\mu_2 = 400$. The calculation of the Between component of Gini index is then based on the "between" Pen's parade :

$$\mathrm{B} = (x_{11}, x_{12}, x_{13}, x_{21}, x_{22}, x_{23}) = (\mu_1, \mu_1, \mu_1, \mu_2, \mu_2, \mu_2) = (200, 200, 200, 400, 400, 400)$$

in which (i) each household is assigned the average of its reference group and (ii) households are ranked on the basis of the values of the group averages. The result is $i_b = 0.167$ as detailed in Table 3.5.

(3) **Calculation of the Within component** Returining to Table 3.5, one reintroduces individual heterogeneity while maintaining the ranking on the average value of the subgroups. The calculation of what constitutes a Quasi- Gini index gives $i_{lex} = 0.278$ (see Table 3.6 for details of the calculation). The point is that the difference between this affordability deficit concentration index (in relation to the lexico-graphical order) and the affordability deficit concentration index in relation to the order on the subgroup averages, that is :

$$i_b - i_{lex} = \frac{5}{18} - \frac{1}{6} = \frac{2}{18} = \frac{1}{9}$$

is precisely equal to the intra component of the Gini index for which we have :

$$f_1 = f_2 = \frac{1}{2}$$

$$\alpha_1 = \frac{X_1}{X} = \frac{600}{1800} = \frac{1}{3}$$

$$\alpha_2 = \frac{X_2}{X} = \frac{1200}{1800} = \frac{2}{3}$$

$$i_w = f_1\alpha_1 \times i_1 + f_2\alpha_2 \times i_2 = \left(\frac{1}{2} \times \frac{1}{3}\right) \times \frac{2}{9} + \left(\frac{1}{2} \times \frac{2}{3}\right) \times \frac{2}{9} = \frac{1}{27} + \frac{2}{27} = \frac{3}{27} = \frac{1}{9}$$

[64] Multiplying the values of the variable of interest by λ does not modify the Lorenz curve and, in so doing, does not modify the area of concentration, nor the double of this area of concentration, what corresponds to the Gini index. This property is interpreted as stating that the inequality mesure depends only on relative unaffordability $x_i/\bar{x}$, $i = 1, \cdots, n$, and, accordingly, will not change whether the affordability deficit is measured in euros or euro cents.

i	x_i	f_i	F_i	α_i	A_i	A_{i-1}	$f_i \times (A_i + A_{i-1})$
1'	200	1/3	1/3	1/6	1/6	0	$\frac{1}{3}\times(0+\frac{1}{6})=\frac{1}{18}$
2'	400	1/3	2/3	2/6	3/6	1/6	$\frac{1}{3}\times(\frac{3}{6}+\frac{1}{6})=\frac{2}{9}$
3'	600	1/3	1	½	1	3/6	$\frac{1}{3}\times(1+\frac{3}{6})=\frac{1}{2}$
∑	1200	1					$\frac{7}{9}$
	$\overline{x}_2=\frac{1200}{3}=400$						$i_2=1-\frac{7}{9}=\frac{2}{9}$

Table 3.4

i	x_i	f_i	F_i	α_i	A_i	A_{i-1}	$f_i \times (A_i + A_{i-1})$
1	200	1/6	1/6	2/18	2/18	0	$\frac{1}{6}\times(\frac{2}{18}+0)$
2	200	1/6	2/6	2/18	4/18	2/18	$\frac{1}{6}\times(\frac{4}{18}+\frac{2}{18})$
3	200	1/6	3/6	2/18	6/18	4/18	$\frac{1}{6}\times(\frac{6}{18}+\frac{4}{18})$
1'	400	1/6	4/6	4/18	10/18	6/18	$\frac{1}{6}\times(\frac{10}{18}+\frac{6}{18})$
2'	400	1/6	5/6	4/18	14/18	10/18	$\frac{1}{6}\times(\frac{14}{18}+\frac{10}{18})$
3'	400	1/6	1	4/18	1	14/18	$\frac{1}{6}\times(1+\frac{14}{18})$
∑	1800	1		1			$\frac{5}{6}$
	$\overline{x}=\frac{1800}{6}=300$						$i_b=1-\frac{5}{6}=\frac{1}{6}=0.167$

Table 3.5

i	x_i	f_i	F_i	α_i	A_i	A_{i-1}	$f_i \times (A_i + A_{i-1})$
1	100	1/6	1/6	1/18	1/18	0	$\frac{1}{6}\times(\frac{1}{18}+0)$
2	200	1/6	2/6	2/18	3/18	1/18	$\frac{1}{6}\times(\frac{3}{18}+\frac{1}{18})$
3	300	1/6	3/6	3/18	6/18	3/18	$\frac{1}{6}\times(\frac{6}{18}+\frac{3}{18})$
1'	200	1/6	4/6	2/18	8/18	6/18	$\frac{1}{6}\times(\frac{8}{18}+\frac{6}{18})$
2'	400	1/6	5/6	4/18	12/18	8/18	$\frac{1}{6}\times(\frac{12}{18}+\frac{8}{18})$
3'	600	1/6	1	6/18	1	12/18	$\frac{1}{6}\times(1+\frac{12}{18})$
∑	1800	1		1			$\frac{13}{18}$
	$\overline{x}=\frac{1800}{6}=300$						$i_{lex}=1-\frac{13}{18}=\frac{5}{18}=0.278$

Table 3.6

(4) Returning to the table above, it appears that the re-ranking of households according the value of their personal unaffordability brings back to Table 3.3, and the value of the Gini index. In this respect, the difference:

$$i-i_{lex}=\frac{8}{27}-\frac{5}{18}=\frac{16-15}{54}=\frac{1}{54}$$

corresponds to the Transvariation component, that is to the contribution (to the value of the Gini index for the population as a whole) generated by the overlap of the distribution of affordability deficit of group 1 on the distribution of affordability deficit of group 2, with the affordability deficit borne by individual 3 in group 1 greater than the one borne by individual 1' in group 2, in a context where affordability deficit of group 1 is, on average, lower than the one of group 2.

A3.3 Additional analysis - Decomposition of Dagum [1997]

A) Measuring inequalities between two groups - the inter-group Gini index The problem here is to measure inequalities in affordability deficit between the members of two groups, for instance group 1 and group 2, using the inter-group block of the matrix of inter-individual gaps:

	$2:1$	$2:2$	$\dots$	$2:n_2$
$1:1$	$\lvert x_{11}-x_{21}\rvert$	$\lvert x_{11}-x_{22}\rvert$	$\dots$	$\lvert x_{11}-x_{2n_2}\rvert$
$1:2$				$\lvert x_{12}-x_{2n_2}\rvert$
$\vdots$	$\vdots$	$\vdots$		$\vdots$
$1:n_1$	$\lvert x_{1n_1}-x_{21}\rvert$	$\lvert x_{1n_1}-x_{22}\rvert$	$\dots$	$\lvert x_{1n_1}-x_{2n_2}\rvert$

Table 3.7

For these purposes, one uses the inter-group Gini index defined as:

$$i_{12}=\frac{\Delta_{12}}{\mu_1+\mu_2} \tag{14.41}$$

with:

$$\Delta_{12}=\frac{1}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{i'=1}^{n_2}\lvert x_{1i}-x_{2i'}\rvert \tag{14.42}$$

the average of the inter-individual inter-group gap (here, of affordability deficit).

Example Returning th the data of the previous numerical example, the block of the inter-individual gaps between the members of group 1 and the members of group 2 is written as:

	$x_{21}=200$	$x_{22}=400$	$x_{23}=600$
$x_{11}=100$	$\lvert 100-200\rvert=100$	$\lvert 100-400\rvert=300$	$\lvert 100-600\rvert=500$
$x_{12}=200$	$\lvert 200-200\rvert=0$	$\lvert 200-400\rvert=200$	$\lvert 200-600\rvert=400$
$x_{13}=300$	$\lvert 300-200\rvert=100$	$\lvert 300-400\rvert=100$	$\lvert 300-600\rvert=300$

The calculation of the inter-group Gini index gives $\mu_1=200$, $\mu_2=400$, and ;

$$\Delta_{12}=\frac{100+300+500+0+200+400+100+100+300}{3\times 3}=\frac{2000}{9}$$

$$i_{12}=\frac{\Delta_{12}}{\mu_1+\mu_2}=\frac{1}{200+400}\times\frac{2000}{9}=\frac{10}{27}=0.370$$

Justification of the measure One starts from the identity :

$$|x_1 + x_2| = x_1 + x_2 - 2\min[x_1, x_2]$$

The sum of the inter-individual inter-group gaps can then be rewritten as :

$$\begin{aligned}\sum_{i=1}^{n_1}\sum_{i'=1}^{n_2}|x_{1i} - x_{2i'}| &= \sum_{i=1}^{n_1}\sum_{i'=1}^{n_2}\left(x_{1i} + x_{2i'} - 2\min[x_{1i}, x_{2i'}]\right) \\ &= \sum_{i=1}^{n_1}\sum_{i'=1}^{n_2}x_{1i} + \sum_{i=1}^{n_1}\sum_{i'=1}^{n_2}x_{2i'} - 2\sum_{i=1}^{n_1}\sum_{i'=1}^{n_2}\min[x_{1i}, x_{2i'}] \\ &= \sum_{i=1}^{n_1}x_{1i}\sum_{i'=1}^{n_2}1 + \sum_{i=1}^{n_1}n_2\mu_2 - 2\sum_{i=1}^{n_1}\sum_{i'=1}^{n_2}\min[x_{1i}, x_{2i'}] \\ &= n_1\mu_1 \times n_2 + n_1 \times n_2\mu_2 - 2\sum_{i=1}^{n_1}\sum_{i'=1}^{n_2}\min[x_{1i}, x_{2i'}] \\ &= n_1 n_2 \times (\mu_1 + \mu_2) - 2\sum_{i=1}^{n_1}\sum_{i'=1}^{n_2}\min[x_{1i}, x_{2i'}]\end{aligned}$$

Accoridngly, the average of the inter-individual inter-group gaps verifies:

$$\Delta_{12} = \frac{1}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{i'=1}^{n_2}|x_{1i} - x_{2i'}| = \mu_1 + \mu_2 - \frac{2}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{i'=1}^{n_2}\min[x_{1i}, x_{2i'}]$$

and we have:

$$\frac{\Delta_{12}}{\mu_1 + \mu_2} = 1 - \frac{2}{n_1 n_2 \times (\mu_1 + \mu_2)}\sum_{i=1}^{n_1}\sum_{i'=1}^{n_2}\min[x_{1i}, x_{2i'}]$$

This statistic (which corresponds to the definition of the inter-group Gini index given above) takes its maximum value, equal to unity, when "1 person has everything and the others have nothing" because, in this case $\min[x_{1i}, x_{2i'}] = 0$ for all i and i'.

B) Economic distance (and overlap ratio) In the following, it is considered that $\mu_2 > \mu_1$ and one separates the average inter-individual inter-group gap Δ_{12} into its two components that are:

(1) Δ_{12}^{+} : the proportion (of the average inter-individual inter-group gap) corresponding to unaffordability pairs $(x_{1i}, x_{2i'})$ for which $x_{2i'} > x_{1i}$ (and that conform to the trend / the order on the averages $\mu_2 > \mu_1$)

(2) Δ_{12}^{-} : the proportion (of the average inter-individual inter-group gap) corresponding to unaffordability pairs $(x_{1i}, x_{2i'})$ for which $x_{1i} > x_{2i'}$ (which is at variance with the trend / the order on the averages $\mu_2 > \mu_1$).

Next, it is defined:

(3) Economic distance

$$d_{12} = \frac{\Delta_{12}^{+} - \Delta_{12}^{-}}{\Delta_{12}} = \frac{\Delta_{12}^{+} - \Delta_{12}^{-}}{\Delta_{12}^{+} + \Delta_{12}^{-}} \quad (14.43)$$

that varies from 0 (in the polar case $\mu_1 = \mu_2$ for which $\Delta_{12}^{-} = \Delta_{12}^{+} = \Delta_{12} / 2$) to 1 (with therefore $\Delta_{12}^{+} = \Delta_{12}$ and $\Delta_{12}^{-} = 0$).

(4) the overlap ratio :

$$\gamma_{12} = 1 - d_{12} = 1 - \frac{\Delta_{12}^{+} - \Delta_{12}^{-}}{\Delta_{12}} = 1 - \frac{\Delta_{12}^{+} - \Delta_{12}^{-}}{\Delta_{12}^{-} + \Delta_{12}^{+}} = \frac{2\Delta_{12}^{-}}{\Delta_{12}} = \frac{\Delta_{12}^{-}}{\Delta_{12} / 2} \quad (14.44)$$

with $\Delta_{12} / 2$ the maximum value of Δ_{12}^{-}[65].

The final point lies in the following property :

Property For the "Between" Distribution for which:

$$i_{12}^{b} = \frac{\mu_2 - \mu_1}{\mu_1 + \mu_2} \quad (14.45)$$

we have:

$$d_{12} \times i_{12} = i_{12}^{b} \quad (14.46)$$

$$i_{12}^{t} = i_{12} - i_{12}^{b} = (1 - d_{12}) \times i_{12} = \gamma_{12} \times i_{12}$$
(14.47).

Numerical example Returning to the data form the numerical example for which $\mu_2 = 400 > \mu_1 = 200$, one breaks down the average inter-individual inter-group gap :

$$\Delta_{12} = \frac{100 + 300 + 500 + 0 + 200 + 400 + 100 + 100 + 300}{3 \times 3} = \frac{2000}{9}$$

Into its two components:

$$\Delta_{12}^{-} = \frac{0 + 100}{3 \times 3} = \frac{100}{9}$$

$$\Delta_{12}^{+} = \frac{100 + 300 + 500 + 200 + 400 + 100 + 300}{3 \times 3} = \frac{1900}{9}$$

The economic distance is then equal to :

[65] We have indeed $0 \le \Delta_{12}^{-} < \Delta_{12} / 2 < \Delta_{12}^{+} \le \Delta_{12} = \Delta_{12}^{-} + \Delta_{12}^{+}$ when $\mu_2 > \mu_1$ and $\Delta_{12}^{-} = \Delta_{12}^{+} = \Delta_{12} / 2$ when $\mu_1 = \mu_2$.

$$d_{12} = \frac{\Delta_{12}^{+} - \Delta_{12}^{-}}{\Delta_{12}} = \frac{1900-100}{2000} = 0.9$$

and the overlap ratio is equal to:

$$\gamma_{12} = 1 - d_{12} = \frac{2\Delta_{12}^{-}}{\Delta_{12}} = \frac{2\times 100}{2000} = 0.1$$

At the same time, for the "Between" distribution, we have :

	$x_{21} = \mu_2 = 400$	$x_{22} = \mu_2 = 400$	$x_{23} = \mu_2 = 400$
$x_{11} = \mu_1 = 200$	$\lvert\mu_1 - \mu_2\rvert = 200$	$\lvert\mu_1 - \mu_2\rvert = 200$	$\lvert\mu_1 - \mu_2\rvert = 200$
$x_{12} = \mu_1 = 200$	$\lvert\mu_1 - \mu_2\rvert = 200$	$\lvert\mu_1 - \mu_2\rvert = 200$	$\lvert\mu_1 - \mu_2\rvert = 200$
$x_{13} = \mu_1 = 200$	$\lvert\mu_1 - \mu_2\rvert = 200$	$\lvert\mu_1 - \mu_2\rvert = 200$	$\lvert\mu_1 - \mu_2\rvert = 200$

with an inter-group Gini index that sets to:

$$i_{12} = \frac{1}{\mu_1 + \mu_2} \times \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{i'=1}^{n_2} \left| x_{1i} - x_{2i'} \right| = \frac{1}{\mu_1 + \mu_2} \times \frac{9\times(\mu_2 - \mu_1)}{9} = \frac{\mu_2 - \mu_1}{\mu_1 + \mu_2} = \frac{400-200}{400+200} = \frac{1}{3}$$

This value has to be compared with the one which is obtained by doing:

$$i_{12}^{b} = d_{12} \times i_{12} = 0.9 \times \frac{10}{27} = \frac{1}{3}$$

Thus, Economic Distance appears to measure the proportion of the inter-group Gini index that is due to the Between component and, by identification, one can deduce that the balance by verifying ;

$$i_{12} - i_{12}^{b} = i_{12} - d_{12} i_{12} = (1 - d_{12}) \times i_{12} = \gamma_{12} \times d_{12}$$

does indeed correspond to the proportion of the inter-group Gini index that is due to transvariation.

C) Dagum decomposition - the case $k = 2$ Returning to table 3.7, with still $\mu_2 > \mu_1$, the latter is partitioned into 4 blocs :

$$\begin{pmatrix} \boxed{B_{11}} & \boxed{B_{12}} \\ \boxed{B_{21}} & \boxed{B_{22}} \end{pmatrix}$$

with:

$\boxed{B_{11}}$ the matrix of the inter-individual gap within the group 1,

$\boxed{B_{12}}$ the matrix of the inter-individual gaps between members of group 1 (in rows) and members of group 2 (in columns),

$\boxed{B_{21}}$ the matrix of inter-individual gaps between members of group 2 (in rows) and members of group 1 (in columns),

$\boxed{B_{22}}$ the matrix of inter-individual gap within the group 2.

Noting I_{11}, I_{12}, I_{21} and I_{22} the sums of the squares of the inter-individual gaps (here, in terms of affordability deficit) within each of the 4 blocks listed above, the Gini index calculated on the population as a whole :

$$i = \frac{\Delta}{2\overline{x}} = \frac{1}{2\overline{x}} \times \frac{1}{n^2} \sum_{i=1}^{n} \sum_{i'=1}^{n} \left| x_i - x_{i'} \right|$$

can be rewritten as:

$$\begin{aligned} i &= \frac{1}{2\overline{x}} \times \frac{I_{11} + I_{12} + I_{21} + I_{22}}{n^2} = \frac{I_{11}}{2n^2\overline{x}} + \frac{I_{22}}{2n^2\overline{x}} + \frac{2I_{12}}{2n^2\overline{x}} \\ &= \frac{n_1^2\mu_1}{n^2\overline{x}} \frac{I_{11}}{2 \times n_1^2\mu_1} + \frac{n_2^2\mu_2}{n^2\overline{x}} \frac{I_{22}}{2 \times n_2^2\mu_2} + \frac{n_1 n_2(\mu_1 + \mu_2)}{n^2\mu} \frac{I_{12}}{n_1 n_2(\mu_1 + \mu_2)} \\ &= \frac{n_1}{n} \frac{n_1\mu_1}{n\overline{x}} \times i_{11} + \frac{n_2}{n_2} \frac{n_2\mu_2}{n\overline{x}} \times i_{22} + \left(\frac{n_1 n_2 \mu_1}{n^2\mu} + \frac{n_1 n_2 \mu_2}{n^2\mu} \right) \times i_{12} \\ &= f_1 \frac{X_1}{X} \times i_{11} + f_2 \frac{X_2}{X} \times i_{12} + \left(\frac{n_2}{n} \frac{n_1\mu_1}{n\mu} + \frac{n_1}{n} \frac{n_2\mu_2}{n\mu} \right) \times i_{12} \\ &= f_1\alpha_1 \times i_{11} + f_2\alpha_2 \times i_{12} + \left(f_2\alpha_1 + f_1\alpha_2 \right) \times i_{12} \end{aligned} \quad (14.48)$$

or in matrix form:

$$i = (f_1, f_2) \begin{pmatrix} i_{11} & i_{12} \\ i_{21} & i_{22} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \quad (14.49)$$

with $i_{12} = i_{21}$ the intergroup Gini index 1-2[66]. In this 2-groups case, the non-within component:

$$i - i_w = \left(f_1\alpha_2 + f_2\alpha_1 \right) \times i_{12} \quad (14.50)$$

is broken down into its two components, Between and Transvariation, making use of the economic distance d_{12} and the overlap ratio $\gamma_{12} = 1 - d_{12}$ as follows:

[66] The notation i_{12} indicates that the inter-group Gini index is calculated with group 1 in rows and group 2 in columns; the notation i_{21} indicates that the inter-group Gini index is calculated with group 2 in rows and group 1 in columns.

$$
\begin{aligned}
i - i_w &= (f_1\alpha_2 + f_2\alpha_1) \times i_{12} = (f_1\alpha_2 + f_2\alpha_1) \times (d_{12} i_{12} + \gamma_{12} i_{12}) \\
&= (f_1\alpha_2 + f_2\alpha_1) \times d_{12} i_{12} + (f_1\alpha_2 + f_2\alpha_1) \times \gamma_{12} i_{12} \\
&= (f_1\alpha_2 + f_2\alpha_1) \times i_{12}^b + (f_1\alpha_2 + f_2\alpha_1) \times i_{12}^t
\end{aligned} \tag{14.51}
$$

with in addition:

$$
i_{12}^b = (\mu_2 - \mu_1) / (\mu_1 + \mu_2)
$$

as stated above. This enables to identify (i) the Between component of the Gini index i (for the Population P as a whole) as given by:

$$
i_b = (f_1\alpha_2 + f_2\alpha_1) \times d_{12} i_{12} = (f_1\alpha_2 + f_2\alpha_1) \times i_{12}^b \tag{14.52}
$$

and (ii) the Transvariation component (of this same Gini index i calculated for the Population P as a whole) as given by:

$$
i_t = (f_1\alpha_2 + f_2\alpha_1) \times \gamma_{12} i_{12} = (f_1\alpha_2 + f_2\alpha_1) \times i_{12}^t \tag{14.53}
$$

with $i_{12}^t = i_{12} - i_{12}^b$.

D) Dagum decomposition - the case $k = 3$ One generalises without additional difficulty to the case with k groups. Considering, by way of illustration and without loss of generality, the case $k = 3$ with $\mu_1 < \mu_2 < \mu_3$, we have (1) the (symmetric) Gini index matrix that is given by:

$$
\begin{pmatrix} i_{11} & i_{12} & i_{13} \\ i_{21} & i_{22} & i_{23} \\ i_{31} & i_{32} & i_{33} \end{pmatrix} \tag{14.54}
$$

and (2) a Gini index calculated for the whole population P that verifies:

$$
i = (f_1, f_2, f_3) \begin{pmatrix} i_{11} & i_{12} & i_{13} \\ i_{21} & i_{22} & i_{23} \\ i_{31} & i_{32} & i_{33} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \tag{14.55}
$$

The Within component of the Gini index i (calculated for the population P) is equal to:

$$
i_w = f_1\alpha_1 \times i_1 + f_2\alpha_2 \times i_2 + f_3\alpha_3 \times i_3 \tag{14.56}
$$

The non-within component $i - i_w$ breaks down into (3) a Between component that sets to:

$$
i_b = (f_1\alpha_2 + f_2\alpha_1) \times i_{12}^b + (f_1\alpha_3 + f_3\alpha_1) \times i_{13}^b + (f_2\alpha_3 + f_3\alpha_2) \times i_{23}^b \tag{14.57}
$$

with:

$$i_{12}^{b}=\frac{\mu_2-\mu_1}{\mu_1+\mu_2},\ i_{13}^{b}=\frac{\mu_3-\mu_1}{\mu_1+\mu_3},\ i_{23}^{b}=\frac{\mu_3-\mu_2}{\mu_2+\mu_3} \tag{14.58}$$

the Between components of the inter-group Gini index i_{12}, i_{13} et i_{23} and (4) a Transvariation component that sets to:

$$i_t=(f_1\alpha_2+f_2\alpha_1)\times i_{12}^{t}+(f_1\alpha_3+f_3\alpha_1)\times i_{13}^{t}+(f_2\alpha_3+f_3\alpha_2)\times i_{23}^{t} \tag{14.59}$$

with:

$$i_{12}^{t}=i_{12}-i_{12}^{b},\ i_{13}^{t}=i_{13}-i_{13}^{b},\ i_{23}^{t}=i_{23}-i_{23}^{b} \tag{14.60}$$

The transvariation components of the inter-group Gini index i_{12}, i_{13} et i_{23}.

Annex 4: Binary classification theory: key metrics

4.1 Basic indicators

Based on the information provided by Table 21, page 97, the tool calculates the values taken by 4 key indicators (and their complements): Sensitivity, Specificity, Positive Predictive Value (or Precision) and Negative Predictive Value.

(1) The first indicator is the **TPR** (True Positive Rate):

$$S_e=\frac{n_{++}}{|\mathrm{D}_+|}=\frac{\underline{Q}_+}{\underline{Q}}=\frac{\underline{Q}_+}{\underline{Q}_++\underline{Q}_-} \tag{14.61}$$

which measures the proportion of overall basic consumption that is correctly subsidised/treated by the tariff (considered by the user). This quantity also gives an estimate of the (conditional) probability that a basic unit (drawn at random from the population of basic units) will be correctly subsidised:

$$\Pr\left[\mathrm{T}_+|\mathrm{D}_+\right]=\frac{\underline{Q}_+}{\underline{Q}} \tag{14.62}$$

The TPR/sensitivity measures the quantity of items to be found that is actually found by the classifier. As such, it is a measure of the exhaustiveness (the property of an enumeration that is found to be complete) of the tariff. The complement of sensitivity (also calculated) is the **FNR** (False Negative Rate) or **Miss Rate**:

$$\mathrm{FNR}=\frac{\underline{Q}_-}{\underline{Q}}=\frac{\underline{Q}_-}{\underline{Q}_++\underline{Q}_-}=1-S_e \tag{14.63}$$

that is the proportion of overall basic consumption that is not subsidised (volume exclusion error). This is also known as the **silence of classifier** (proportion of relevant items not detected / not processed).

(2) The second indicator is **TNR** (True Negative Rate):

$$S_p = \frac{n_{--}}{|\mathrm{D}_-|} = \frac{\left(Q-\underline{Q}\right)_-}{Q-\underline{Q}} = \frac{\left(Q-\underline{Q}\right)_-}{\left(Q-\underline{Q}\right)_- + \left(Q-\underline{Q}\right)_+} \tag{14.64}$$

defined as the proportion of overall non basic consumption that is correctly treated (and therefore not subsidised) by the tariff (considered by the user). This quantity gives an estimate of the (conditional) probability that a non-basic unit (drawn at random from the population of service units that do not meet basic needs) will rightly not be subsidised:

$$\Pr\left[\mathrm{T}_- | \mathrm{D}_-\right] = \frac{\left(Q-\underline{Q}\right)_-}{Q-\underline{Q}} \tag{14.65}$$

The complement to the TNR is the **FPR** (False Positive Rate):

$$\bar{S}_p = 1 - S_p = \frac{n_{+-}}{|\mathrm{D}_-|} = \frac{\left(Q-\underline{Q}\right)_+}{Q-\underline{Q}} = \frac{\left(Q-\underline{Q}\right)_+}{\left(Q-\underline{Q}\right)_- + \left(Q-\underline{Q}\right)_+} \tag{14.66}$$

(also calculated and displayed) that gives an estimate of the conditional probability that a non-basic unit (drawn at random from the population of service units that do not meet basic needs) will be wrongly subsidised:

$$\Pr\left[\mathrm{T}_+ | \mathrm{D}_-\right] = \frac{\left(Q-\underline{Q}\right)_+}{Q-\underline{Q}} \tag{14.67}$$

(3) Indicators 3 and 4 **are the Positive Predictive Value**:

$$\mathrm{PPV} = \frac{n_{++}}{|\mathrm{T}_+|} = \frac{\underline{Q}_+}{Q_+} \tag{14.68}$$

(or **Accuracy**) which gives the probability that a subsidised unit (drawn at random from the population of units that are subsidised by the tariff) is a basic unit and the **Negative Predicted Value**:

$$\mathrm{NPV} = \frac{n_{--}}{|\mathrm{T}_-|} = \frac{\left(Q-\underline{Q}\right)_-}{Q_-} \tag{14.69}$$

which gives the probability that a non-subsidised unit (drawn at random from the population of units that are not subsidised by the tariff) is not a basic unit / is indeed a non-basic unit. The complements to the PPV and NPV (also calculated and displayed) are, respectively, the **False Discovery Rate**:

$$\mathrm{FDR} = 1 - \mathrm{PPV} = \frac{\left(Q - \underline{Q}\right)_{+}}{Q_{+}} \tag{14.70}$$

which gives the probability that a subsidised unit (drawn at random from the population of subsidised units) is wrongly subsidised, i.e. the percentage of subsidised consumption that should not be subsidised/which is for non-basic uses and the **False Omission Rate**:

$$\mathrm{FOR} = 1 - \mathrm{NPV} = \frac{\underline{Q}_{-}}{Q_{-}} \tag{14.71}$$

which gives the probability that a non-subsidised unit (drawn at random from the population of non-subsidised units) is a basic unit, i.e. the percentage of non-subsidised consumption that should be subsidised.

(4) The calculation of these (8) indicators is then completed by the calculation of prevalence:

$$\text{Prévalence} = p = \frac{\left|\mathrm{D}_{+}\right|}{n} = \frac{\underline{Q}}{Q} \tag{14.72}$$

(based on the information provided by the econometric model of household water demand, after reprocessing by the user) since the quality of targeting as measured by PPV (mutatis mutandis by NPV) proves to be all the better the higher the prevalence (ceteris paribus).

Details The PPV (Accuracy) by satisfying the relation:

$$\mathrm{PPV} = \frac{\underline{Q}_{+}}{Q_{+}} = \frac{S_e \times p}{S_e \times p + \left(1 - S_p\right)\left(1 - p\right)} \tag{14.73}$$

increases with prevalence *p*:

$$\frac{\partial \mathrm{PPV}}{\partial p} = \frac{\partial}{\partial p}\left[\frac{\underline{Q}_{+}}{Q_{+}}\right] = \frac{S_e \times \left(1 - S_p\right)}{\left[\ldots\right]^2} > 0 \tag{14.74}$$

(except in special cases $S_e = 0$ and/or $S_p = 1$) while the NPV by satisfying:

$$\mathrm{NPV} = \frac{\left(Q - \underline{Q}\right)_{-}}{Q_{-}} = \frac{S_p \times (1 - p)}{S_p \times (1 - p) + \left(1 - S_e\right) \times p} \tag{14.75}$$

decreases with prevalence *p*:

$$\frac{\partial \mathrm{NPV}}{\partial p} = \frac{\partial}{\partial p}\left[\frac{\left(Q - \underline{Q}\right)_{-}}{Q_{-}}\right] = -\frac{\left(1 - S_e\right) \times S_p}{\left[\ldots\right]^2} < 0 \tag{14.76}$$

(except in special cases $S_e = 1$ and/or $S_p = 0$).

4.2 The ROC Space

It is here to assess the quality of the tariff and, in particular, its ability to discriminate between basic and non-basic units. Different approaches are possible (and implemented by the tool) with the first that focus on the Sensitivity / Accuracy (PPV) pair which is in itself a natural (but non-scalar) measure of the effectiveness of tariff (which is considered by the user) as a classifier. In particular, (1) a value of the Sensitivity:

$$S_e = \frac{\underline{Q}_+}{\underline{Q}} = \frac{\underline{Q}_+}{\underline{Q}_+ + \underline{Q}_-} = \frac{1}{1 + \frac{\underline{Q}_-}{\underline{Q}_+}} \tag{14.77}$$

that tends towards 1 means that the tariff tends to subsidise all basic units (with a few false negatives $\underline{Q}_-$ relative to $\underline{Q}_+$) while (2) a value of the Accuracy:

$$\mathrm{PPV} = \frac{\underline{Q}_+}{Q_+} = \frac{\underline{Q}_+}{\underline{Q}_+ + \left(Q - \underline{Q}\right)_+} = \frac{1}{1 + \frac{\left(Q - \underline{Q}\right)_+}{\underline{Q}_+}} \tag{14.78}$$

that tends towards 1 means that the tariff tends to subsidise only basic units (with a few false positives $\left(Q - \underline{Q}\right)_+$ relative to $\underline{Q}_+$). In this way, a situation in which $S_e = 1$ and $\mathrm{PPV} = 1$ describes an ideal situation in which the tariff subsidises all basic units ($S_e = 1$; no exclusion error) and only basic units ($\mathrm{PPV} = 1$; no inclusion error).

To position the tariff in relation to this ideal situation (and also in relation to 3 other polar cases; see below), the tool calculates at first 3 other indicators that are:

(1) the "Positive likelihood ratio" LR_+,

(2) the "Negative likelihood ratio' LR_-,

(3) the Diagnostic odds ratio DOR ,

and next displays a diagram called the ROC[67] Space showing all of this information and, also, enabling to compare the performance of different IBTs (as classifiers).

Diagram Components The ROC space is constructed by plotting (1) Anti-Specificity on the x-axis:

$$\bar{S}_p = 1 - S_p = \frac{\left(Q - \underline{Q}\right)_+}{Q - \underline{Q}} \tag{14.79}$$

therefore, the propensity of the tariff to subsidise a unit of consumption (cubic metre) which is a non-basic one, and (2) Sensitivity on the y-axis:

$$S_e = \frac{\underline{Q}_+}{\underline{Q}} \tag{14.80}$$

therefore, the propensity of the tariff to subsidise a unit of consumption (cubic metre) which is a basic one. In addition to the point:

---

[67] Receiver Operating Characteristic.

$$\mathrm{T} = \left(1 - S_p, S_e\right) = \left(\frac{\left(Q - \underline{Q}\right)_+}{Q - \underline{Q}}, \frac{\underline{Q}_+}{\underline{Q}} \right) \tag{14.81}$$

(with coordinates $\left(\frac{9}{100}, \frac{2}{3}\right)$ on Figure 14, page 98) that corresponds to the characteristics of the IBT which is evaluated/tested by the user, the diagram (that fits into the unit square) is also showing some additional constitutive components with:

- the 4 vertices $\mathrm{O} = (0,0)$, $\mathrm{A} = (1,0)$, $\mathrm{B} = (1,1)$ and $\mathrm{C} = (1,0)$ (of the unit square),
- the 45° line,
- the point $\mathrm{P} = (p, p)$ (with coordinates (0.6,0.6) on Figure 14) on the 45°line with (as a reminder) $p = \underline{Q}/Q$ the prevalence / the share of basic consumption $\underline{Q}$ in overall consumption Q,
- the lines OT (in red) and BT (in green).

Finally, the point $S = (s_T, s_T)$ (not shown in Figure 14) with $s_T = Q_+/Q$ the subsidy rate is also displayed (for the visualization of an adjusted J of Youden that is discussed later in the text).

Chart interpretation The vertices of the unit square represent 4 main pricing schemes, which are used as reference points.

(1) The origin:

$$\mathrm{O} = \left(\bar{S}_p, S_e\right) = \left(1 - S_p, S_e\right) = \left(\frac{\left(Q - \underline{Q}\right)_+}{Q - \underline{Q}}, \frac{\underline{Q}_+}{\underline{Q}} \right) = (0,0) \tag{14.82}$$

describes the performance of a tariff that never subsidises. This case includes, as a special case, the TBSE $(F, \pi) = \left(\frac{CF}{n}, c\right)$ which does not subsidise any unit, whether it is basic or non-basic. At this point, the PPV is not defined since:

$$\mathrm{PPV} = \frac{\underline{Q}_+}{Q_+} = \frac{0}{0} \tag{14.83}$$

("indeterminate form") while the NPV sets to:

$$\mathrm{NPV} = \frac{\left(Q - \underline{Q}\right)_-}{Q_-} = \frac{Q - \underline{Q}}{Q} = 1 - \frac{\underline{Q}}{Q} = 1 - p \tag{14.84}$$

(2) The vertice:

$$\mathrm{A} = \left(\bar{S}_p, S_e\right) = \left(1 - S_p, S_e\right) = \left(\frac{\left(Q - \underline{Q}\right)_+}{Q - \underline{Q}}, \frac{\underline{Q}_+}{\underline{Q}} \right) = (0,1) \tag{14.85}$$

accounts for an ideal classifier that never makes a mistake with an inclusion error that sets to zero with:

$$\bar{S}_p = 1 - S_p = \frac{\left(Q - \underline{Q}\right)_+}{Q - \underline{Q}} = 0 \Leftrightarrow \left(Q - \underline{Q}\right)_+ = 0 \tag{14.86}$$

and an exclusion error that sets also to zero with:

$$S_e = \frac{\underline{Q}_+}{\underline{Q}} = \frac{\underline{Q}_+}{\underline{Q}_+ + \underline{Q}_-} = 1 \Leftrightarrow \underline{Q}_- = 0 \tag{14.87}$$

This case can be interpreted as describing a situation in which each household *i* would be offered a personalised tariff, with subsidised consumption blocks covering precisely its basic consumption $\underline{q}_i$. In this case, all the subsidised units are basic units, with a PPV equal to 1:

$$\text{PPV} = \frac{\underline{Q}_+}{Q_+} = \frac{\underline{Q}_+}{\underline{Q}_+ + \left(Q-\underline{Q}\right)_+} = \frac{\underline{Q}_+}{\underline{Q}_+ + 0} = 1 \tag{14.88}$$

and all units that are not subsidised are non-basic units with an NPV also equal to 1:

$$\text{NPV} = \frac{\left(Q-\underline{Q}\right)_-}{Q_-} = \frac{\left(Q-\underline{Q}\right)_-}{\left(Q-\underline{Q}\right)_- + \underline{Q}_-} = \frac{Q-\underline{Q}}{Q-\underline{Q}+0} = 1 \tag{14.89}$$

(3) The vertice:

$$\text{B} = \left(\overline{S}_p, S_e\right) = \left(1-S_p, S_e\right) = \left(\frac{\left(Q-\underline{Q}\right)_+}{Q-\underline{Q}}, \frac{\underline{Q}_+}{\underline{Q}}\right) = (1,1) \tag{14.90}$$

describes a situation in which all consumption units are subsidised with:

$$\left(Q-\underline{Q}\right)_+ = Q-\underline{Q}$$

$$\underline{Q}_+ = \underline{Q}$$

and therefore:

$$Q_+ = \underline{Q}_+ + \left(Q-\underline{Q}\right)_+ = \underline{Q} + Q - \underline{Q} = Q.$$

This scenario reflects a policy in which the tariff subsidises all the units consumed, with for instance $\pi_1 < \pi_2 < \pi_3 < \pi_4 < c$ in the case of an IBT4, and taxes the access fee to guarantee (as far as possible) the financial equilibrium of the service. Exclusion errors are then at their minimum and inclusion errors at their maximum (for a given overall consumption/service level Q). In this case, the PPV is equal to the prevalence *p*:

$$\text{PPV} = \frac{\underline{Q}_+}{Q_+} = \frac{\underline{Q}}{Q} = p \tag{14.91}$$

and the NPV is undefined:

$$\text{VPN} = \frac{\left(Q-\underline{Q}\right)_-}{Q_-} = \frac{0}{0} \tag{14.92}$$

("indeterminate form").

(4) The vertice:

$$\text{C} = \left(\overline{S}_p, S_e\right) = \left(1-S_p, S_e\right) = \left(\frac{\left(Q-\underline{Q}\right)_+}{Q-\underline{Q}}, \frac{\underline{Q}_+}{\underline{Q}}\right) = (1,0) \tag{14.93}$$

reports a situation in which the classifier is always wrong with:

$$\left(Q-\underline{Q}\right)_+ = Q-\underline{Q}$$

$$\underline{Q}_+ = 0$$

This case describes a system of personalised degressive tariffs in which non-basic consumption would be subsidised and basic consumption would be taxed (to guarantee, as far as possible, the financial equilibrium of the service). This scenario is therefore outside the scope of the tool.

Besides these 4 reference points, the diagram is also showing:

(5) the point $\mathrm{P} = (p, p)$, with $p = \underline{Q}/Q$ the prevalence rate, which has a PPV equal to the prevalence rate $p = \underline{Q}/Q$ and a NPV equal to its complement, $1 - p = \left(Q - \underline{Q}\right)/Q$ [68].

The point P describes the expected effects of a random pricing policy that allocates subsidies at random, deciding for each unit of service with probability $p = \underline{Q}/Q$ that it is subsidised, and with probability $1 - p$ that it is not (and, therefore, taxed as soon as $\pi_j \neq c, \forall j = 1, \ldots, p$). In this respect, point P constitutes a reference point from which the quality of IBT targeting can be assessed, by comparing the latter with what would be obtained if this random policy were carried out[69]. In this way, the added value linked to the deterministic nature of the pricing policy is identified, at least as far as the quality of targeting is concerned[70]. This type of comparison is often used in the literature on the evaluation of public policies, in particular to measure the quality of targeting of aid programmes in favour of poor households (with the calculation of Ω ratios and the pivot value $\Omega = 1$ (see section 8)).

(6) the radius OT which represents the iso-PPV line of level $\mathrm{PPV}_0 = \mathrm{PPV}_T$ with slope:

$$a_+ = \frac{S_e(T)}{1-S_p(T)} = \frac{\underline{Q}_+}{\underline{Q}} \times \frac{1}{\dfrac{\left(Q-\underline{Q}\right)_+}{Q-\underline{Q}}} = \frac{\Pr\left[T_+ \middle| D_+\right]}{\Pr\left[T_+ \middle| D_-\right]}$$

also equal to the odds ratio:

$$a_+ = \frac{\mathrm{PPV}}{1-\mathrm{PPV}} \times \frac{1}{\dfrac{p}{1-p}} = \frac{\Pr\left[D_+ \middle| T_+\right]}{\Pr\left[D_- \middle| T_+\right]} \times \frac{1}{\dfrac{p}{1-p}} = \mathrm{LR}_+ \times \frac{1}{\dfrac{p}{1-p}}$$

with:

68 More generally, all the points located on the 45° line have a PPV equal to the prevalence rate $p = \underline{Q}/Q$ and an NPV equal to its complement, $1 - p$ (as it can be shown by evaluating (12.24) at $1 - S_p = S_e$, and (12.26) at $S_p = 1 - S_e$,

69 By the law of large numbers, the subsidy rate of this random pricing policy can be considered to be equal to the prevalence rate *p*.

70 Modulo the adjustment of the subsidy rate $s^+ = Q^+/Q$ which is not necessarily equal to the prevalence rate $p = \underline{Q}/Q$ for the IBT that is considered/tested by the user. We return to this point later.

$$\mathrm{LR}_+ = \frac{\Pr\left[D_+|T_+\right]}{\Pr\left[D_-|T_+\right]} = \frac{\Pr\left[D_+|T_+\right]}{1-\Pr\left[D_+|T_+\right]} = \frac{\underline{Q}_+}{Q_+} \times \frac{1}{\frac{\left(Q-\underline{Q}\right)_+}{Q^+}} = \frac{\underline{Q}_+}{Q_+} \times \frac{1}{1-\frac{\underline{Q}_+}{Q_+}} = \frac{\underline{Q}_+}{\left(Q-\underline{Q}\right)_+} \qquad (14.94)$$

the "Positive likelihood ratio" (for a non-basic unit that is subsidised, how many basic units are subsidised), the value of which is also displayed by the tool;

(7) the BT line which represents the iso-VPN line of level $\mathrm{NPV}_0 = \mathrm{NPV}_T$ with calculation of the slope:

$$a_- = \frac{1-S_e(T)}{S_p(T)} = \ldots = \frac{1-\mathrm{VPN}}{\mathrm{VPN}} \times \frac{1}{\frac{p}{1-p}} = \frac{\Pr\left[D_+|T_-\right]}{\Pr\left[D_-|T_-\right]} \times \frac{1}{\frac{p}{1-p}} = \mathrm{LR}_- \times \frac{1}{\frac{p}{1-p}}$$

with:

$$\mathrm{LR}_- = \frac{\Pr\left[D_+|T_-\right]}{\Pr\left[D_-|T_-\right]} = \frac{\underline{Q}_-}{Q_-} \times \frac{1}{\frac{\left(Q-\underline{Q}\right)_-}{Q_-}} = \frac{\underline{Q}_-}{\left(Q-\underline{Q}\right)_-} \qquad (14.95)$$

the "Negative likelihood ratio" (how many unsubsidised basic units for one unsubsidised non basic unit, that is a service unit meeting comfort or luxury needs), the value of which is also displayed by the tool.

Next, since the tariff is all the better (*i*) the smaller the Anti-Specificity:

$$x_{\mathrm{T}} = \bar{S}_p(T) = 1 - S_p(T) = \frac{\left(Q-\underline{Q}\right)_+}{Q-\underline{Q}}$$

("not many subsidised units in units that should not be subsidised") and (*ii*) the greater the Sensitivity:

$$y_{\mathrm{T}} = S_e(T) = \frac{\underline{Q}_+}{\underline{Q}} = \frac{\underline{Q}_+}{\underline{Q}_+ + \underline{Q}_-}$$

("not many unsubsidised units in units that should be subsidised"), so (*iii*) the closer the point $T = \left(1 - S_p(T), S_e(T)\right)$ (that accounts for the IBT tested by the user) to vertice $A = (1,0)$ (that accounts for the ideal personalised pricing system described above), it is also shown that, in terms of targeting, the IBT will also be all the more effective:

- the greater the value by which the fractional odds $p/(1-p)$ has to be multiplied in the population of the subsidised units (in particular, the slope coefficient a_+ and the LR_+ are expected to be greater than unity),
- the smaller the value by which the fractional odds $p/(1-p)$ has to be multiplied in the population of the non-subsidised units (in particular, the slope coefficient a_- and the ratio LR_- are expected to be less than unity),

starting from point $\mathrm{P} = (p, p)$ and a random allocation of subsidised status with a probability $p = \underline{Q}/Q$ (see above). The discriminatory power of pricing is then measured by the Diagnostic Odds Ratio:

$$\mathrm{DOR} = \frac{\mathrm{LR}_+}{\mathrm{LR}_-} = \ldots = \frac{\underline{Q}_+}{\underline{Q}_-} \times \frac{\left(Q-\underline{Q}\right)_-}{\left(Q-\underline{Q}\right)_+} \tag{14.96}$$

(whose value is also displayed by the tool) with:

- $\underline{Q}_+/\left(Q-\underline{Q}\right)_+$ a ratio indicating the number of basic units that are rightly subsidised for one non-basic unit that is wrongly subsidised,
- $\underline{Q}_-/\left(Q-\underline{Q}\right)_-$ a ratio indicating the number of basic units that are wrongly not-subsidised for one non-basic unit that is rightly not subsidised.

The value obtained for the DOR is assessed compared to the value of 1, that is the value taken by the indicator with a random allocation of subsidised status with probability $p = \underline{Q}/Q$.

4.3 Other indicators

The tool concludes this assessment with the calculation of 7 other indicators to gauge the quality of the IBT (as a classifier): Accuracy and Adjusted ACC, Youden's J and adjusted Youden's J, Kappa Score, Cohen's Ratio, and Jaccard Index.

(1) Accuracy is the proportion of instances (here, cubic metres) correctly classified:

$$\mathrm{ACC} = \frac{n_{++} + n_{--}}{n} = \frac{\underline{Q}_+ + \left(Q-\underline{Q}\right)_-}{Q} \tag{14.97}$$

The authors (systematically) point out that the (often high) values taken by this indicator need to be put into perspective. In this case, however, a natural reference point is the TBSE for which $\underline{Q}_+ = 0$ and $\left(Q-\underline{Q}\right)_- = Q-\underline{Q}$ are available. To measure the performance (in terms of targeting) of an IBT, this leads to an **adjusted ACC** defined as:

$$\mathrm{ACC}^* = \mathrm{ACC} - \mathrm{ACC}_{\mathrm{TBSE}} = \frac{\underline{Q}_+ + \left(Q-\underline{Q}\right)_-}{Q} - \frac{0 + Q - \underline{Q}}{Q} = \frac{\underline{Q}_+}{Q} - \frac{\left(Q-\underline{Q}\right)_+}{Q} \tag{14.98}$$

that is the difference between the aggregate inclusion gain and the aggregate inclusion error (as a percentage of the service level).

(2) **Youden's J** consists in calculating the quantity:

$$\begin{aligned} J &= S_e + S_p - 1 = S_e - \left(1 - S_p\right) = S_e - \bar{S}_p = \frac{\underline{Q}_+}{\underline{Q}} - \frac{\left(Q-\underline{Q}\right)_+}{Q-\underline{Q}} \\ &= \left(S_e - p\right) - \left(\bar{S}_p - p\right) = \left(S_e - p\right) + \left(p - \bar{S}_p\right) \end{aligned} \tag{14.99}$$

that is the Manhattan distance between:

(i) the point $\mathrm{T} = \left(S_e(\mathrm{T}), \bar{S}_p(\mathrm{T})\right) = \left(S_e(\mathrm{T}), 1 - S_p(\mathrm{T})\right)$ which is characteristic of the IBT tested by the user with a subsidy rate set to $s_+ = Q_+/Q$, *(ii)* the point $P = (p, p)$ which is characteristic of the random pricing policy consisting of allocating subsidised status, among the population of the Q cubic metres that are consumed, with probability $p = \underline{Q}/Q$.

It should be noted that the Youden's J is also equal to the Manhattan distance from the point:

$$S = \left(s_T, s_T\right) = \left(\frac{Q_+}{Q}, \frac{Q_+}{Q}\right)$$

insofar as:

$$J^* = \left(S_e - s_T^+\right) - \left(\bar{S}_p - s_T^+\right) = \left(S_e - s_T^+\right) + \left(s_T^+ - \bar{S}_p\right) = \left(\frac{\underline{Q}_+}{\underline{Q}} - \frac{Q_+}{Q}\right) + \left(\frac{Q_+}{Q} - \frac{\left(Q - \underline{Q}\right)_+}{Q - \underline{Q}}\right) \qquad (14.100)$$

(3) **Kappa score** Originally, the indicator noted I_κ measures the degree of agreement between 2 assessors ("inter-rater reliability") who have to qualify (yes / no) each application in a set of n applications. Once the selection process implemented, the data to be analysed is a contingency table of the form:

		Appraiser B		
		Yes	No	
Appraiser A	Yes	n_{11}	n_{12}	$n_{1.}$
	No	n_{21}	n_{22}	$n_{2.}$
		$n_{.1}$	$n_{.2}$	n

Table 72 : Contingency table - calculation of Kappa score

with A the identity of the first assessor, B that of the second). Kappa score is then calculated as:

$$I_\kappa = \frac{p_0 - p^*}{1 - p^*} \qquad (14.101)$$

with:

$$p_0 = \frac{n_{11} + n_{22}}{n} \qquad (14.102)$$

the observed frequency of agreement (concordance) and p^* the frequency of agreement (concordance) that would be observed under the hypothesis of independence, that is:

$$p^* = \frac{n_{1.}}{n} \times \frac{n_{.1}}{n} + \frac{n_{2.}}{n} \times \frac{n_{.2}}{n} \qquad (14.103)$$

The value obtained is interpreted as expressing "the extent to which the value of the agreement observed among the assessors goes beyond chance", i.e. exceeds what would have been obtained under the assumption of independence. The term $1 - p^*$ (in the denominator) gives the degree of agreement that is achievable beyond chance, while the term $p_0 - p^*$ (in the numerator) gives the degree of agreement beyond chance that is actually achieved.

The Kappa score verifies the following properties :

(i) The value of the Kappa score is between $-p^*/(1-p^*)$ and 1.

(ii) Polar case $I_\kappa = 1$ establishes complete agreement/systematic concordance between the 2 assessors because in this case:

		Appraiser B		
		Yes	No	
Appraiser A	Yes	n_{11}	0	$n_{1.}$
	No	0	n_{22}	$n_{2.}$
		$n_{.1}$	$n_{.2}$	n

Table 73 : showing the calculation of Kappa score with systematic agreement

with therefore $p_0 = \frac{n_{11}+n_{22}}{n} = \frac{n}{n} = 1$ and $I_\kappa = \frac{p_0-p^*}{1-p^*} = \frac{1-p^*}{1-p^*} = 1$.

(iii) Polar case $I_\kappa = -p^*/(1-p^*)$ establishes total/complete disagreement between the 2 assessors because in this case:

		Appraiser B		
		Yes	No	
Appraiser A	Yes	0	n_{12}	$n_{1.}$
	No	n_{21}	0	$n_{2.}$
		$n_{.1}$	$n_{.2}$	n

Table 74 : showing the calculation of Kappa score with complete mismatch

with therefore $p_0 = \frac{n_{11}+n_{22}}{n} = \frac{0}{n} = 0$ and $I_\kappa = \frac{p_0-p^*}{1-p^*} = \frac{0-p^*}{1-p^*} = -\frac{p^*}{1-p^*}$.

(iv) Refence case $I_\kappa = 0 \Leftrightarrow p_0 = p^*$ indicates that concordance of the rankings can be considered as the result of chance.

Application The tool applies the Kappa score calculation to Table 21, page 97, by having the tariff play the role of evaluator A and Nature the role of evaluator B (this image is common in decision theory to reflect the environment in which a decision-maker operates). The percentage of instances (i.e. units of service) correctly classified (processed):

$$p_0 = \frac{\underline{Q}_+ + \left(Q-\underline{Q}\right)_-}{Q} \tag{14.104}$$

then corresponds to the ACC (see above) and we have:

$$p^* = \frac{\underline{Q}}{Q} \times \frac{Q_+}{Q} + \frac{Q-\underline{Q}}{Q} \times \frac{Q_-}{Q} = p \times s_T^+ + \left(1-p\right) \times \left(1-s_T^+\right) = \frac{\underline{Q}Q_+ + \left(Q-\underline{Q}\right)Q_-}{Q^2} \tag{14.105}$$

$$1-p^* = \frac{\underline{Q}}{Q} \times \frac{Q_-}{Q} + \frac{Q-\underline{Q}}{Q} \times \frac{Q_+}{Q} = p \times \left(1-s_T^+\right) + \left(1-p\right) \times s_T^+ = \frac{\underline{Q}Q_- + \left(Q-\underline{Q}\right)Q_+}{Q^2} \tag{14.106}$$

In fine, Kappa score is equal to:

$$I_\kappa = \frac{p_0 - p^*}{1-p^*} = \ldots = \frac{\left[\underline{Q}_+ + \left(Q-\underline{Q}\right)_-\right]\times Q - \left[\underline{Q}Q_+ + \left(Q-\underline{Q}\right)Q_-\right]}{\underline{Q}Q_- + \left(Q-\underline{Q}\right)Q_+} \quad (14.107)$$

(4) **Cohen's ratio** One difficulty with the Kappa score is that the value of the indicator varies, for the same level of agreement (ACC), with the classification behaviour of assessors A and B, in this case with the prevalence rate $p = \underline{Q}/Q$ and the rate of subsidy of the tariff policy $s_+ = Q_+/Q$.

Details We have:

$$I_\kappa = \frac{p_0 - p^*}{1-p^*} \quad \text{with} \quad p_0 = \frac{\underline{Q}_+ + \left(Q-\underline{Q}\right)_-}{Q} = \text{ACC}$$

(by definition of the Kappa index) with the value of the index that is decreasing in p^* :

$$\frac{\partial I_\kappa}{\partial p^*} = \ldots = -\frac{1-p_0}{\left(1-p^*\right)^2} < 0$$

and a concordance rate under the hypothesis of independence which verifies:

$$p^* = \frac{\underline{Q}}{Q}\times\frac{Q_+}{Q} + \frac{Q-\underline{Q}}{Q}\times\frac{Q_-}{Q} = p\times s_T^+ + (1-p)\times\left(1-s_T^+\right)$$

(as highlighted above). ■.

Applied to the question of the performance of a classifier / the measurement of inclusion and exclusion errors with an IBT, this property is not without its problems. Firstly, it implies that two pricing policies can only be compared, strictly speaking, on the basis of the same prevalence $p = \underline{Q}/Q$, i.e. the same overall consumption Q, and a constant (aggregate) subsidy rate Q_+/Q. Secondly, it means that the maximum value of the Kappa score can no longer be given by 1 when the marginal distributions (scoring behaviour) of the 2 assessors differ, i.e. when the prevalence rate $p = \underline{Q}/Q$ and the subsidy rate $s_+ = Q_+/Q$ are not equal (this point in detailed at the end of this paragraph devoted to the presentation of Cohen's ratio).

An alternative consists in comparing the performance (in terms of classification) of different pricing policies using Cohen's ratio:

$$\frac{I_\kappa}{I_\kappa^{\max}} = \frac{p_0 - p^*}{1-p^*} \times \frac{1}{\dfrac{p_0^{\max} - p^*}{1-p^*}} = \frac{p_0 - p^*}{p_0^{\max} - p^*} \quad (14.108)$$

which consists (simply) in calculating the % of agreement that is actually achieved relative to the maximum agreement that is (actually) possible to achieve given the (fixed) values of $p = \underline{Q}/Q$ (prevalence) and $s_+ = Q_+/Q$ (subsidy). Ultimately, this means applying the following formulas depending on whether the subsidy rate $s_+ = Q_+/Q$ is greater than, equal to or lower than the prevalence rate $p = \underline{Q}/Q$:

Case 1 : $Q_+ > \underline{Q} \Leftrightarrow s > p$ with:

$$p_0^{\max} = \mathrm{ACC}_{\max} = \frac{\underline{Q} + Q^{-}}{Q} = p + 1 - s \tag{14.109}$$

and :

$$\frac{I_\kappa}{I_\kappa^{\max}} = \frac{p_0 - p^*}{p_0^{\max} - p^*} = \frac{\left[\underline{Q}_+ + \left(Q - \underline{Q}\right)_-\right] \times Q - \left[\underline{Q}Q_+ + \left(Q - \underline{Q}\right)Q_-\right]}{\left(\underline{Q} + Q_-\right) \times Q - \left[\underline{Q}Q_+ + \left(Q - \underline{Q}\right)Q_-\right]} \tag{14.110}$$

Case 2 : $Q_+ = \underline{Q} \Leftrightarrow s = p$ with:

$$p_0^{\max} = \mathrm{ACC}_{\max} = \frac{\underline{Q} + Q - \underline{Q}}{Q} = 1 \tag{14.111}$$

and a Cohen ratio that is also equal to the Kappa score:

$$\frac{I_\kappa}{I_\kappa^{\max}} = \frac{p_0 - p^*}{p_0^{\max} - p^*} = \frac{p_0 - p^*}{1 - p^*} = I_\kappa \tag{14.112}$$

- Case 3 : $Q_+ < \underline{Q} \Leftrightarrow s < p$ with:

$$p_0^{\max} = \mathrm{ACC}_{\max} = \frac{Q_+ + Q - \underline{Q}}{Q} = s + 1 - p = 1 - \frac{\underline{Q} - Q_+}{Q} = 1 - \left(p - s\right) = s + 1 - p \tag{14.113}$$

and a Cohen's ratio given by:

$$\frac{I_\kappa}{I_\kappa^{\max}} = \frac{p_0 - p^*}{p_0^{\max} - p^*} = \frac{\left[\underline{Q}_+ + \left(Q - \underline{Q}\right)_-\right] \times Q - \left[\underline{Q}Q_+ + \left(Q - \underline{Q}\right)Q_-\right]}{\left(Q_+ + Q - \underline{Q}\right) \times Q - \left[\underline{Q}Q_+ + \left(Q - \underline{Q}\right)Q_-\right]} \tag{14.114}$$

Details On the maximum value of Kappa index (and calculation of Cohen's ratio) **In case 1**, $Q_+ > \underline{Q}$ (in which case $Q_- < Q - \underline{Q}$) and the "at best-confusion matrix" is given by:

	D_+	D_-	
T_+	$Q^+ = \underline{Q}$	$(Q - \underline{Q})^+ = Q^+ - \underline{Q}$	Q^+
T_-	$\underline{Q}^- = 0$	$(-\underline{Q})^- = Q^-$	Q^-
	$\underline{Q}$	$Q - \underline{Q}$	Q

Table 75 : Calculation of Cohen score - case I

with:

$$p_0^{\max} = \mathrm{ACC}_{\max} = \frac{\underline{Q} + Q_-}{Q} = p + 1 - s \tag{14.115}$$

$$p^* = \frac{\underline{Q}Q_+ + \left(Q - \underline{Q}\right)Q_-}{Q^2} = ps + \left(1 - p\right)\left(1 - s\right) \tag{14.116}$$

$$p_0^{\max} - p^* = \frac{\underline{Q} + Q_-}{Q} - \frac{\underline{Q}Q_+ + \left(Q - \underline{Q}\right)Q_-}{Q^2} \tag{14.117}$$

and a Cohen's ratio that sets to:

$$\frac{I_\kappa}{I_\kappa^{\max}} = \frac{p_0 - p^*}{p_0^{\max} - p^*} = \frac{\left[Q_+ + \left(Q - \underline{Q}\right)_-\right] \times Q - \left[\underline{Q}Q_+ + \left(Q - \underline{Q}\right)Q_-\right]}{\left(\underline{Q} + Q_-\right) \times Q - \left[\underline{Q}Q_+ + \left(Q - \underline{Q}\right)Q_-\right]} \tag{14.118}$$

In case 2, $Q_+ = \underline{Q}$ and the at best-confusion matrix is given by:

	D_+	D_-	
T_+	$\underline{Q}^+ = \underline{Q}$	$(Q - \underline{Q})^+ = 0$	Q^+
T_-	$\underline{Q}^- = 0$	$(Q - \underline{Q})^- = Q - \underline{Q}$	Q^-
	$\underline{Q}$	$Q - \underline{Q}$	Q

Table 76 : Calculation of Cohen score - case II

This gives the ideal classifier for which:

$$p_0^{\max} = \mathrm{ACC}_{\max} = \frac{\underline{Q} + Q - \underline{Q}}{Q} = 1 \tag{14.119}$$

with a Cohen ratio equal to the Kappa score:

$$\frac{I_\kappa}{I_\kappa^{\max}} = \frac{p_0 - p^*}{p_0^{\max} - p^*} = \frac{p_0 - p^*}{1 - p^*} = I_\kappa \tag{14.120}$$

and a value p^* that sets to:

$$p^* = \frac{\underline{Q} \times Q_+}{Q^2} + \frac{\left(Q - \underline{Q}\right) \times Q_-}{Q^2} = \frac{\underline{Q} \times \underline{Q}}{Q^2} + \frac{\left(Q - \underline{Q}\right) \times \left(Q - \underline{Q}\right)}{Q^2} = p^2 + \left(1 - p\right)^2 = s^2 + \left(1 - s\right)^2 \tag{14.121}$$

In case 3, $Q_+ < \underline{Q}$ (in which case $Q_- > Q - \underline{Q}$) and the "at best-matrix confusion matrix" is given by:

	D_+	D_-	
T_+	$\underline{Q}^+ = Q^+$	$(Q - \underline{Q})^+ = 0$	Q^+
T_-	$\underline{Q}^- = \underline{Q} - Q^+$	$(Q - \underline{Q})^- = Q - \underline{Q}$	Q^-
	$\underline{Q}$	$Q - \underline{Q}$	Q

Table 77 : Calculation of Cohen score - case III

with:

$$p_0^{\max} = \mathrm{ACC}_{\max} = \frac{Q_+ + Q - \underline{Q}}{Q} = s + 1 - p = 1 - \frac{\underline{Q} - Q_+}{Q} = 1 - \left(p - s\right) = s + 1 - p \tag{14.122}$$

and a Cohen's ratio that sets to:

$$\frac{I_\kappa}{I_\kappa^{\max}} = \frac{p_0 - p^*}{p_0^{\max} - p^*} = \frac{\left[Q_+ + \left(Q - \underline{Q}\right)_-\right] \times Q - \left[\underline{Q}Q_+ + \left(Q - \underline{Q}\right)Q_-\right]}{\left(Q_+ + Q - \underline{Q}\right) \times Q - \left[\underline{Q}Q_+ + \left(Q - \underline{Q}\right)Q_-\right]} \tag{14.123}$$

■.

(6) **Jaccard Index** The last index is the Jaccard index, which is calculated as follows:

$$\text{JAC} = \frac{\left|\underline{Q} \cap S_+\right|}{\left|\underline{Q} \cup S_+\right|} = \frac{\underline{Q}_+}{\underline{Q}_+ + \underline{Q}_- + \left(Q - \underline{Q}\right)_+}$$

This index measures the similarity between the two sets constituted by (1) the population of cubic metres (service units) meeting basic needs and (2) the population of subsidised cubic metres (service units). This statistic, giving the proportion of elements common to both sets, varies from 0 (in this case, the number of basic units that are subsidised $\underline{Q}_+$ is equal to 0) to 1 (in this case, the number of basic units that are not subsidised $\underline{Q}_-$ and the number of subsidised units that do not meet basic needs $\left(Q - \underline{Q}\right)_+$ are both equal to 0).

Annex 5: Setting the relative distribution curve

As a reminder, the net subsidy received on its water consumption by household *i* is given by:

$$\left(t_{iq} + s_{iq}\right)_- = -\min\left[t_{iq} + s_{iq}, 0\right] \tag{14.124}$$

where s_{iq} is the gross subsidy (counted negatively) and t_{iq} is the gross contribution to service funding through the consumption q. The mass of these net subsidies implemented by the tariff is then equal to:

$$\left(T_q + S_q\right)_- = \sum_{i=1}^{n}\left(t_{iq} + s_{iq}\right)_- \tag{14.125}$$

On this basis, the following operations are performed:

- Households are ordered according to the values of the standard of living, from the lowest value to the highest (households are then indiced by i^* in this sorted series) with:

$$F_{i^*} = \frac{i^*}{n} \tag{14.126}$$

the standardised rank of household *i* in the Pen parade of living standards.

- The cumulative increasing frequencies of the mass of the AFE net subsidy are computed with:

$$A_h = \sum_{i^*=1}^{h} \frac{\left(t_{i^*q} + s_{i^*q}\right)}{\left(T_q + S_q\right)_-} \tag{14.127}$$

for $h = 1 \ldots, n$.

Then, in the plane (F, A) (within the unit square) :

- the $n+1$ points with coordinates:

$$(0,0),\left(\frac{1}{n},A_1\right),\left(\frac{2}{n},A_2\right),\ldots,\left(\frac{n-1}{n},A_{n-1}\right),\left(\frac{n}{n},A_n\right)=(1,1)$$

are represented, next connected by line segments to obtain an estimate of the concentration curve, noted $A = A(F)$, of the AFE net subsidies in relation to the standard of living.

- the 45° line and the poverty line $F = F_{\text{Poor}}$ (given the threshold value for the standard of living entered by the user in the Social Data section of the General Data tab) are displayed.

By construction, the concentration curve $A = A(F)$ thus obtained gives the percentage of the mass of the subsidies that goes to $100F\%$ the least wealthy members of the population, for all values $F \in [0,1]$. The subsidy system is then qualified as:

- redistributive for sure when the concentration curve is above the 45° line for any value $0 < F < 1$.

In this case, the least wealthy $100F\%$ of the population receive a share of the total subsidies that is greater than $100F\%$ and this observation holds for all values of the standardised rank $F \in]0,1[$, and:

- anti-redistributive for sure when the concentration curve lies below the 45° line for any value $0 < F < 1$.

In this case, the least wealthy $100F\%$ of the population receive a share of the total subsidies that is less than $100F\%$, whatever the value of $F \in]0,1[$.

In these configurations,

- the degree of redistribution/anti-redistribution can be assessed by computing the equivalent of a Gini index, geometrically the (oriented) area bounded by the 45° line $A = F$ that depicts a situation in which the mass of subsidies would be distributed equally across the population (what applies to TBSE since, with the TBSE, each household receives an equal share of the mass of subsidies that sets to 0) and the concentration curve $A = A(F)$

The latter varies from -1 (in this case, the concentration curve is almost vertical at $F = 0$ with all subsidies going to the poorest household) to 1 (in this case, the concentration curve is almost vertical at $F = 1$ with all subsidies going to the richest household).

The Omega ratio is read with the ratio of two (vertical) segments: the one corresponding to the value returned by the $A(F)$ function at $F = F_{\text{Poor}}$, $A(F_{\text{Poor}})$ vs. the one corresponding to the value returned by the 45° line at $F = F_{\text{Poor}}$, that is F_{Poor}.

As shown in Figure 55-5, graphing the relative benefit distribution curve is of particular interest when the curve intersects the 45° line at 1 point, or even at several points (the system represented in Figure 55-5 appears to be redistributive for households in poverty but anti-redistributive for households in extreme poverty).

Figure 55 : Relative Benefit Distribution Curves with Progressive and Regressive Distributions

(Ω ratio values and Quasi-Gini Index (visualization))

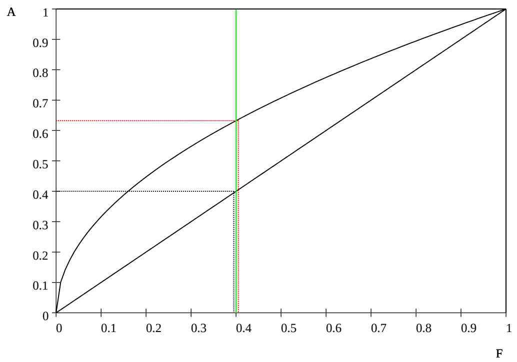


fig 55.1: Redistribution system

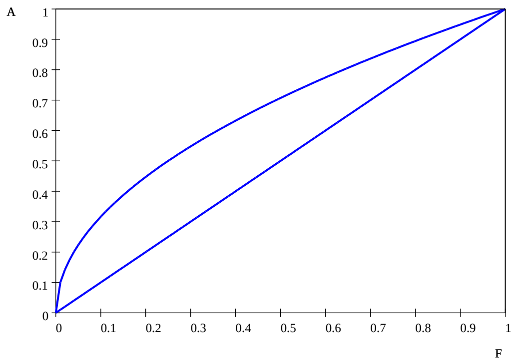


fig 55.2: Negative Quasi-Gini Index

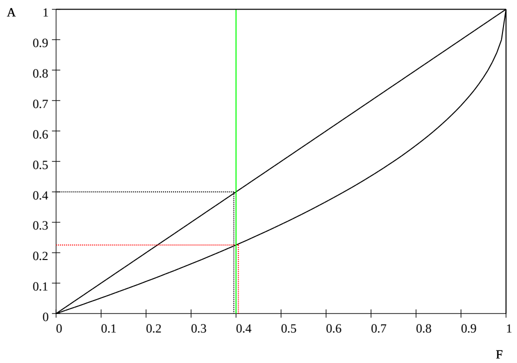


fig 55.3: Anti-redistribution system

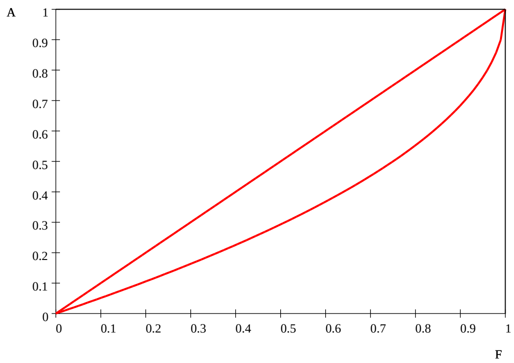


fig 55.4: Positive Quasi-Gini Index

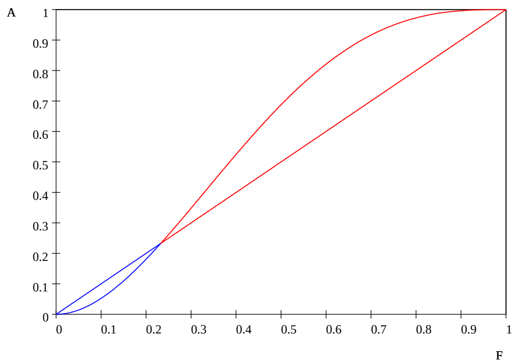


fig 55.5: Non-dominance

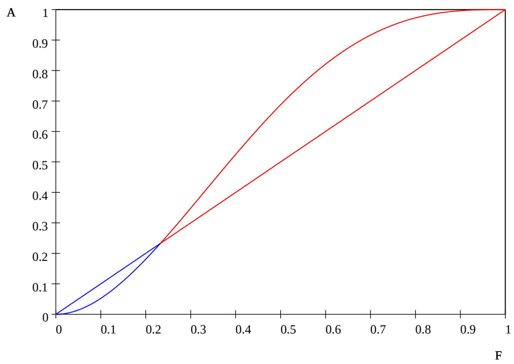


fig 55.6: Positive Quasi-Gini Index

INN WATER

Annex 6: Ex post tariff adjustment – calculation of the catching-up / rebate rate

As a reminder, the equation to be solved (in x) is given by:

$$\sum_{i=1}^{n}\left[F-D_i+(\pi_i+x)q_i^{\mathrm{IBT}}\right]=CF+c\times\sum_{i=1}^{n}q_i^{\mathrm{IBT}} \tag{14.128}$$

The member on the left can be rewritten as:

$$\sum_{i=1}^{n}\left[F-D_i+(\pi_i+x)q_i^{\mathrm{IBT}}\right]=\sum_{i=1}^{n}\left[F-D_i+\pi_i q_i^{\mathrm{IBT}}\right]+\sum_{i=1}^{n}xq_i^{\mathrm{IBT}}=\sum_{i=1}^{n}\left[F-D_i+\pi_i q_i^{\mathrm{IBT}}\right]+x\times n\overline{q}_i^{\mathrm{IBT}} \tag{14.129}$$

with $\overline{q}_i^{\mathrm{IBT}}$ the average household consumption for the IBT tariff, which is evaluated/tested by the user. Then, we have:

$$\begin{aligned} x\times n\overline{q}_i^{\mathrm{IBT}} &= CF+c\sum_{i=1}^{n}q_i^{\mathrm{IBT}}-\sum_{i=1}^{n}\left[F-D_i+\pi_i q_i^{\mathrm{IBT}}\right]=CF-nF-\sum_{i=1}^{n}\left[(\pi_i-c)q_i^{\mathrm{IBT}}-D_i\right] \\ &=-n\times\left(F-\frac{CF}{n}\right)-\sum_{i=1}^{n}\left[(\pi_i-c)q_i^{\mathrm{IBT}}-D_i\right]=-n\times c_{i0}-\sum_{i=1}^{n}\left(t_{iq}+s_{i,q}\right) \end{aligned} \tag{14.130}$$

and finally :

$$\begin{aligned} x &= \frac{1}{\overline{q}_i^{\mathrm{IBT}}}\times\left[\left(F-\frac{CF}{n}\right)-\frac{1}{n}\sum_{i=1}^{n}\left[(\pi_i-c)q_i^{\mathrm{IBT}}-D_i\right]\right]=-\frac{1}{\overline{q}_i^{\mathrm{IBT}}}\times\left[c_{i0}+\frac{1}{n}\sum_{i=1}^{n}\left(t_{iq}+s_{i,q}\right)\right] \\ &=-\frac{c_{i0}+\overline{t_q+s_q}}{\overline{q}_i^{\mathrm{IBT}}} \end{aligned} \tag{14.131}$$

with $c_{i0}=F-\frac{CF}{n}$ the subsidy / taxation on the access fee and:

$$\overline{t_q+s_q}=\frac{1}{n}\sum_{i=1}^{n}\left(t_{iq}+s_{i,q}\right) \tag{14.132}$$

the average net contribution, through their consumption, of subscribers (households) to the financing of the system.

Annexe 7 Factor Decomposition of Gini index

A) Initial elements The starting point is household income, which is assumed to be supplemented by *q* sources of income:

$$R_i=x_i^1+x_i^2+\ldots+x_i^q=\sum_{m=1}^{q}x_i^m$$

with R_i the income level of household *i*, *i* varying from 1 to *n*, and x_i^m the level of its income of type *m*, $m=1,\ldots,q$ (the higher indices represent factors, the lower indices individuals). Without loss of generality, it will be assumed in the following that households are ranked according to the value of their income, from the smallest one to the largest one with :

$R_1 \leq R_2 \leq \cdots \leq R_n$

Next, it is considered:

(1) the Pen's parade of household income $\mathbf{R} = (R_1, R_2, \dots, R_n)$,

(2) $i(\mathbf{R})$: the Gini index of household income distribution $\mathbf{R} = (R_1, R_2, \dots, R_n)$,

(3) $F_\mathbf{R} = \left(\frac{1}{n}, \frac{2}{n}, \dots, 1\right)$ the series of normalised ranks giving the (normalised) ranking of household *i*, *i* varying from 1 to *n*, in the Pen's parade of incomes $\mathbf{R} = (R_1, R_2, \dots, R_n)$,

(4) $\mathbf{x}^m = (x_1^m, x_2^m, \dots, x_n^m)$ the series of the income source *m* associated with the Pen's parade of household income $\mathbf{R} = (R_1, R_2, \dots, R_n)$ (in the series $\mathbf{x}^m$, households are ranked according to the value of their total income R_i),

(5) $C_m(F)$ the concentration curve for revenue source *m*, x^m, in relation to revenue R :

$$\frac{1}{n} \to C\left(\frac{1}{n}\right) = \frac{x_1^m}{x_1^m + x_2^m + \dots + x_n^m}$$

$$\frac{2}{n} \to C\left(\frac{2}{n}\right) = \frac{x_1^m + x_2^m}{x_1^m + x_2^m + \dots + x_n^m}$$

$$\vdots$$

$$\frac{n}{n} \to C(1) = \frac{x_1^m + x_2^m + \dots + x_n^m}{x_1^m + x_2^m + \dots + x_n^m} = 1$$

therefore, the graph of the function that gives the share of total income from source *m*, $X^m = x_1^m + x_2^m + \cdots + x_n^m$, that benefits to the $100\frac{i}{n}$% of the population of households with incomes below R_i, $i = 1, \dots, n$.

(6) $c(\mathbf{x}^m, \mathbf{R})$ the concentration index (or coefficient) of income source *m*, x^m, in relation to income R :

$$c(\mathbf{x}^m, \mathbf{R}) = 2\int_0^1 \left(t - C_m(t)\right) dt$$

Similar to the Gini index, this coefficient corresponds to the ratio of the area of concentration, defined as the area bounded by the concentration curve $C_m(F)$ and the 45° line, to the "maximum" area of concentration, associated with a polar case in which all the income from source *m*, X^m, benefits to individual *i* with the highest income R_i, i.e. the individual with normalised $F_\mathbf{R} = 1$. It is also to keep in mind that, as the concentration curve can lie above the 45° line (unlike a Lorenz curve), a concentration coefficient can take on a negative value (unlike the Gini index)[71].

---

[71] More generally, the concentration curve of source income *m* over income *R* differs from the Lorenz curve of source income *m* in which households are ranked in ascending order of their source income *m*. In this respect, (i) a Lorenz

B) Decompositions Once these elements introduced, decomposition of Rao [1969] states that the Gini index of the income distribution:

$$\mathbf{R} = (R_1, R_2, \ldots, R_n)$$

can be broken down as follows:

$$i_R = \sum_{m=1}^{q} 2 \times \frac{\text{cov}(x^m, F_R)}{\bar{R}} = \sum_{m=1}^{q} \frac{\bar{x}^m}{\bar{R}} \times \frac{2 \times \text{cov}(x^m, F_R)}{\bar{x}^m} = \sum_{m=1}^{q} \theta_m \times c(x^m, R) = \sum_{m=1}^{q} S_m \qquad (14.133)$$

with:

- s_m the contribution of the income source x^m to the Gini index of household income R,

- $\text{cov}(x^m, F_R)$ the covariance between (i) source income *m* and (ii) the normalized rank of households in the Pen's parade of household income $\mathbf{R} = (R_1, R_2, \ldots, R_n)$:

$$\text{cov}(x^m, F_R) = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{i}{n} - \bar{F}_R \right) \left(x_i^m - \bar{x}^m \right),$$

- $\bar{x}_m$ the average of incomes of source *m* x_1^m, x_2^m, ..., x_n^m

- $\bar{R}$ the average of household incomes R_1, R_2, ..., R_n,

- θ_m the share of income of type *m* in total income $R = R_1 + R_2 + \cdots + R_n$:

$$\theta_m = \frac{x_1^m + x_2^m + \ldots + x_n^m}{R_1 + R_2 + \ldots + R_n} = \frac{n \times \bar{x}_m}{n \times \bar{R}} = \frac{\bar{x}_m}{\bar{R}}$$

The decomposition of Pyatt, Chen and Fei [1980] takes Rao's decomposition a step further by revealing the Gini index of the income source *m*, that can be calculated as[72] :

$$i(\mathbf{x}^m) = \frac{2 \times \text{cov}(x^m, F_{x^m})}{\bar{x}^m} \qquad (14.134)$$

with F_{x^m} the series of normalized ranks $\left(\frac{1}{n}, \frac{2}{n}, \ldots, 1\right)$ in the Pen's parade of revenues from source *m* :

curve constitutes a particular concentration curve with, for example, the Lorenz curve of source *m* income which is the concentration curve of source *m* income in relation to source *m* income. Moreover, (ii) a concentration curve, for example of the income source x^m en relation to income *R*, is necessarily located above the Lorenz curve of the income source x^m (as explained in Lambert [2001], "the income share of the poorest 100*p* per cent of the population is equal to or lower than the income share of any other 100*p* per cent of the population" and "difference between the two curves, if any, is accounted for by differences in the rankings of the distributions in question"), page 29.

[72] This formula is a third manner of calculating the Gini index. See Lerman & Yitzhaki [1984] and Yitzhaki [1998]].

$$\mathbf{x}_m^* = \left(x_{m,1}^*, x_{m,2}^*, \ldots, x_{m,n}^*\right),\ x_{m,1}^* \le x_{m,2}^* \le \ldots \le x_{m,n}^*,.$$

Returning to equation (14.133), we get:

$$i_R = \sum_{m=1}^{q} \frac{\bar{x}^m}{\bar{R}} \times \frac{\text{cov}(x^m, F_R)}{\text{cov}(x^m, F_{x^m})} \times \frac{2 \times \text{cov}(x^m, F_{x^m})}{\bar{x}^m} = \sum_{m=1}^{q} \theta_m \times R(x^m, R) \times i(x^m)$$

with:

$$R(x^m, R) = \frac{\text{cov}(x^m, F_R)}{\text{cov}(x^m, F_{x^m})}$$

the so-called " Gini-correlation term ", also equal to the ratio of the concentration coefficient $c(\mathbf{x}^m, \mathbf{R})$ to the Gini index $i(\mathbf{x}^m)$ of source m income,.that takes values in the interval $[-1,1]$. It should be noted that the Gini-correlation term is equal to 1 (-1) if source m income is a monotonically increasing (decreasing) function of total income R, with the values closer to -1 or 1 that indicate a strong relationship between x^m and R.

C) Implementation With regard to the implementation of these breakdowns, the approach is as follows. The income of a household after payment of its IBT bill is given by:

$$R_i^{\text{IBT}} = R_i - T_{\text{IBT}}\left(q_i^{\text{IBT}}\right)$$

and after payment of the TBSE invoice by :

$$R_i^{\text{TBSE}} = R_i - T_{\text{TBSE}}\left(q_i^{\text{TBSE}}\right) = R_i - \left(\frac{CF}{n} + c \times q_i^{\text{TBSE}}\right)$$

Accordingly, we have:

$$R_i^{\text{IBT}} = R_i^{\text{TBSE}} + \left(R_i^{\text{IBT}} - R_i^{\text{TBSE}}\right) = R_i^{\text{TBSE}} + T_{\text{TBSE}}\left(q_i^{\text{TBSE}}\right) - T_{\text{IBT}}\left(q_i^{\text{IBT}}\right)$$

with:

$$\begin{aligned} T_{\text{TBSE}}\left(q_i^{\text{TBSE}}\right) - T_{\text{IBT}}\left(q_i^{\text{IBT}}\right) &= T_{\text{TBSE}}\left(q_i^{\text{TBSE}}\right) - T_{\text{TBSE}}\left(q_i^{\text{IBT}}\right) + T_{\text{TBSE}}\left(q_i^{\text{IBT}}\right) - T_{\text{IBT}}\left(q_i^{\text{IBT}}\right) \\ &= \left[\left(\frac{CF}{n} + cq_i^{\text{TBSE}}\right) - \left(\frac{CF}{n} + cq_i^{\text{IBT}}\right)\right] + T_{\text{TBSE}}\left(q_i^{\text{IBT}}\right) - T_{\text{IBT}}\left(q_i^{\text{IBT}}\right) \\ &= -c \times \left(q_i^{\text{IBT}} - q_i^{\text{TBSE}}\right) + T_{\text{TBSE}}\left(q_i^{\text{IBT}}\right) - T_{\text{IBT}}\left(q_i^{\text{IBT}}\right) \end{aligned}$$

(one recognizes the incentive effect) and :

$$
\begin{aligned}
T_{\text{TBSE}}\left(q_i^{\text{IBT}}\right)-T_{\text{IBT}}\left(q_i^{\text{IBT}}\right) &= \frac{CF}{n}+cq_i^{\text{IBT}}-\left(F+\pi_1 k_1+\pi_2 q_i^{\text{IBT}}-k_1\right) \\
&= \frac{CF}{n}+ck_1+c\left(q_i^{\text{IBT}}-k_1\right)-\left(F+\pi_1 k_1+\pi_2 q_i^{\text{IBT}}-k_1\right) \\
&= \frac{CF}{n}-F+\left(c-\pi_1\right)k_1+\left(c-\pi_2\right)\left(q_i^{\text{IBT}}-k_1\right) \\
&= -\left(F-\frac{CF}{n}\right)-\left(\pi_1-c\right)k_1-\left(\pi_2-c\right)\left(q_i^{\text{IBT}}-k_1\right)
\end{aligned}
$$

by considering, to simplify the presentation, a household located in block 2 (without loss of generality). Next, the following truncated variables are introduced:

- Taxation (Margin) on Access Fee:

$$
c_{i0}^{+}=\max\left[0,F-\frac{CF}{n}\right]=\begin{cases} F-\dfrac{CF}{n} & if \quad F\geq\dfrac{CF}{n} \\ 0 & if \quad F<\dfrac{CF}{n} \end{cases}
$$

- Subsidy on Access Fee:

$$
c_{i0}^{-}=-\min\left[0,F-\frac{CF}{n}\right]=\begin{cases} \dfrac{CF}{n}-F & if \quad F\leq\dfrac{CF}{n} \\ 0 & if \quad F>\dfrac{CF}{n} \end{cases}
$$

- Taxation (margin) on Block 1 consumption:

$$
c_{i1}^{+}=\max\left[0,\left(\pi_1-c\right)k_1\right]=\begin{cases} \left(\pi_1-c\right)k_1 & if \quad \pi_1\geq c \\ 0 & if \quad \pi_1<c \end{cases}
$$

- Subsidy on Block 1 consumption:

$$
c_{i1}^{-}=-\min\left[0,\left(\pi_1-c\right)k_1\right]=\begin{cases} \left(c-\pi_1\right)k_1 & if \quad c\geq\pi_1 \\ 0 & if \quad c<\pi_1 \end{cases}
$$

- Taxation (margin) on Block 2 consumption:

$$c_{i2}^{+} = \max\left[0, (\pi_2 - c)\left(q_i^{\text{IBT}} - k_1\right)\right] = \begin{cases} (\pi_2 - c)\left(q_i^{\text{IBT}} - k_1\right) & if \quad \pi_2 \geq c \\ 0 & if \quad \pi_2 < c \end{cases}$$

- Subsidy on Block 2 consumption:

$$c_{i2}^{-} = -\min\left[0, (\pi_2 - c)\left(q_i^{\text{IBT}} - k_1\right)\right] = \begin{cases} (c - \pi_2)\left(q_i^{\text{IBT}} - k_1\right) & if \quad c \geq \pi_2 \\ 0 & if \quad c < \pi_2 \end{cases}$$

(this notation because subsidies are outgoing flows for the Operator and taxes (margins) are incoming flows). Next, we have:

$$\begin{aligned} T_{\text{TBSE}}\left(q_i^{\text{IBT}}\right) - T_{\text{IBT}}\left(q_i^{\text{IBT}}\right) &= -\left(F - \frac{CF}{n}\right) - (\pi_1 - c)k_1 - (\pi_2 - c)\left(q_i^{\text{IBT}} - k_1\right) \\ &= c_{i0}^{-} - c_{i0}^{+} + c_{i1}^{-} - c_{i1}^{+} + c_{i2}^{-} - c_{i2}^{+} \\ &= c_{i0}^{-} + \left(-c_{i0}^{+}\right) + c_{i1}^{-} + \left(-c_{i1}^{+}\right) + c_{i2}^{-} + \left(-c_{i2}^{+}\right) \end{aligned}$$

and:

$$\begin{aligned} R_i^{\text{IBT}} &= R_i^{\text{TBSE}} + T_{\text{TBSE}}\left(q_i^{\text{TBSE}}\right) - T_{\text{IBT}}\left(q_i^{\text{IBT}}\right) \\ &= R_i^{\text{TBSE}} - c \times \left(q_i^{\text{IBT}} - q_i^{\text{TBSE}}\right) + T_{\text{TBSE}}\left(q_i^{\text{IBT}}\right) - T_{\text{IBT}}\left(q_i^{\text{IBT}}\right) \\ &= R_i^{\text{TBSE}} + c \times \left(q_i^{\text{TBSE}} - q_i^{\text{IBT}}\right) + c_{i0}^{-} + \left(-c_{i0}^{+}\right) + c_{i1}^{-} + \left(-c_{i1}^{+}\right) + c_{i2}^{-} + \left(-c_{i2}^{+}\right) \end{aligned}$$

Denoting:

$$\text{IC}_i^{-} = \max\left[0, c \times \left(q_i^{\text{TBSE}} - q_i^{\text{IBT}}\right)\right] = \begin{cases} c \times \left(q_i^{\text{TBSE}} - q_i^{\text{IBT}}\right) & if \quad q_i^{\text{IBT}} \leq q_i^{\text{TBSE}} \\ 0 & if \quad q_i^{\text{IBT}} > q_i^{\text{TBSE}} \end{cases}$$

$$\text{IC}_i^{+} = \max\left[0, c \times \left(q_i^{\text{IBT}} - q_i^{\text{TBSE}}\right)\right] = \begin{cases} c \times \left(q_i^{\text{IBT}} - q_i^{\text{TBSE}}\right) & if \quad q_i^{\text{IBT}} \geq q_i^{\text{TBSE}} \\ 0 & if \quad q_i^{\text{IBT}} < q_i^{\text{TBSE}} \end{cases}$$

the truncated variables for the incentive effect and the "adverse incentive effect", we have ultimately:

$$\begin{aligned} R_i^{\mathrm{IBT}} &= R_i^{\mathrm{TBSE}} + c\times\left(q_i^{\mathrm{TBSE}} - q_i^{\mathrm{IBT}}\right) + c_{i0}^{-} + \left(-c_{i0}^{+}\right) + c_{i1}^{-} + \left(-c_{i1}^{+}\right) + c_{i2}^{-} + \left(-c_{i2}^{+}\right) \\ &= R_i^{\mathrm{TBSE}} + \mathrm{IC}_i^{-} - \mathrm{IC}_i^{+} + c_{i0}^{-} + \left(-c_{i0}^{+}\right) + c_{i1}^{-} + \left(-c_{i1}^{+}\right) + c_{i2}^{-} + \left(-c_{i2}^{+}\right) \\ &= R_i^{\mathrm{TBSE}} + \mathrm{IC}_i^{-} + \left(-\mathrm{IC}_i^{+}\right) + c_{i0}^{-} + \left(-c_{i0}^{+}\right) + c_{i1}^{-} + \left(-c_{i1}^{+}\right) + c_{i2}^{-} + \left(-c_{i2}^{+}\right) \end{aligned}$$

This relationship enables the Gini index to be broken down by factor (the introduction of taxation does not pose any additional difficulties once the amounts of the excise duties and of VAT have been isolated).

Annex 8: Additional disagregations of funding sources for basic service

We have:

(1) As regards the term $\frac{\underline{C}}{C}\times\underline{\gamma}$ relating to (probable) basic support:

Disaggregation by Groups G1 vs. G2 (customer segments) :

$$\frac{C(\underline{Q})}{C(Q)}\times\underline{\gamma} = \frac{\underline{C}}{C}\times\frac{\underline{R}}{\underline{C}} = \frac{\underline{C}}{C}\times\frac{\underline{R}_1+\underline{R}_2}{\underline{C}_1+\underline{C}_2} = \frac{\underline{C}_1}{C}\times\frac{\underline{R}_1}{\underline{C}_1} + \frac{\underline{C}_2}{C}\times\frac{\underline{R}_2}{\underline{C}_2} = \frac{\underline{C}_1}{C}\times\underline{\gamma}_1 + \frac{\underline{C}_2}{C}\times\underline{\gamma}_2$$

Disaggregation by Services (EP vs. A) :

$$\frac{C(\underline{Q})}{C(Q)}\times\underline{\gamma} = \frac{\underline{C}}{C}\times\frac{\underline{R}}{\underline{C}} = \frac{\underline{C}}{C}\times\frac{\underline{R}_{\mathrm{EP}}+\underline{R}_{\mathrm{A}}}{\underline{C}_{\mathrm{EP}}+\underline{C}_{\mathrm{A}}} = \frac{\underline{C}_{\mathrm{EP}}}{C}\times\frac{\underline{R}_{\mathrm{EP}}}{\underline{C}_{\mathrm{EP}}} + \frac{\underline{C}_{\mathrm{A}}}{C}\times\frac{\underline{R}_{\mathrm{A}}}{\underline{C}_{\mathrm{A}}} = \frac{\underline{C}_{\mathrm{EP}}}{C}\times\underline{\gamma}_{\mathrm{EP}} + \frac{\underline{C}_{\mathrm{A}}}{C}\times\underline{\gamma}_{\mathrm{A}}$$

Disaggregation by Groups and Services (EP-G1, EP-G2 and A-G2) :

$$\begin{aligned} \frac{C(\underline{Q})}{C(Q)}\times\underline{\gamma} &= \frac{\underline{C}}{C}\times\frac{\underline{R}}{\underline{C}} = \frac{\underline{C}}{C}\times\frac{\underline{R}_1^{\mathrm{EP}}+\underline{R}_2^{\mathrm{EP}}+\underline{R}_2^{\mathrm{A}}}{\underline{C}_1^{\mathrm{EP}}+\underline{C}_2^{\mathrm{EP}}+\underline{C}_2^{\mathrm{A}}} = \frac{\underline{C}_1^{\mathrm{EP}}}{C}\times\frac{\underline{R}_1^{\mathrm{EP}}}{\underline{C}_1^{\mathrm{EP}}} + \frac{\underline{C}_2^{\mathrm{EP}}}{C}\times\frac{\underline{R}_2^{\mathrm{EP}}}{\underline{C}_2^{\mathrm{EP}}} + \frac{\underline{C}_2^{\mathrm{A}}}{C}\times\frac{\underline{R}_2^{\mathrm{A}}}{\underline{C}_2^{\mathrm{A}}} \\ &= \frac{\underline{C}_1^{\mathrm{EP}}}{C}\times\underline{\gamma}_1^{\mathrm{EP}} + \frac{\underline{C}_2^{\mathrm{EP}}}{C}\times\underline{\gamma}_2^{\mathrm{EP}} + \frac{\underline{C}_2^{\mathrm{A}}}{C}\times\underline{\gamma}_2^{\mathrm{A}} \end{aligned}$$

(2) As regards the term relating to the captive but non-basic consumption:

Disaggregation by Groups G1 vs. G2 (customer segments) :

$$\begin{aligned} \frac{c\left(Q_0-\underline{Q}\right)}{C(Q)}\times\gamma_{\mathrm{Fixe\_HB}} &= \frac{C_0-\underline{C}}{C(Q)}\times\frac{R_0-\underline{R}}{C_0-\underline{C}} = \frac{C_0-\underline{C}}{C(Q)}\times\frac{R_{0,1}+R_{0,2}-\underline{R}_1-\underline{R}_2}{C_{0,1}+C_{0,2}-\underline{C}_1-\underline{C}_2} \\ &= \frac{C_{0,1}-\underline{C}_1}{C(Q)}\times\frac{R_{0,1}-\underline{R}_1}{C_{0,1}-\underline{C}_1} + \frac{C_{0,2}-\underline{C}_2}{C(Q)}\times\frac{R_{0,2}-\underline{R}_2}{C_{0,2}-\underline{C}_2} \\ &= \frac{C_{0,1}-\underline{C}_1}{C(Q)}\times\gamma_{\mathrm{Fixe\_HB,1}} + \frac{C_{0,2}-\underline{C}_2}{C(Q)}\times\gamma_{\mathrm{Fixe\_HB,2}} \end{aligned}$$

Disaggregation by Services (EP vs. A) :

$$
\begin{aligned}
\frac{c\left(Q_0-\underline{Q}\right)}{C\left(Q\right)}\times\gamma_{\text{Fixe\_HB}} &= \frac{c_{\text{EP}}\left(Q_0-\underline{Q}\right)+c_{\text{A}}\left(Q_0^{\text{A}}-\underline{Q}_{\text{A}}\right)}{C\left(Q\right)}\times\gamma_{\text{Fixe\_HB}} \\
&= \frac{C_0-\underline{C}}{C\left(Q\right)}\times\frac{R_0-\underline{R}}{C_0-\underline{C}} = \frac{C_0-\underline{C}}{C\left(Q\right)}\times\frac{R_0^{\text{EP}}+R_0^{\text{A}}-\underline{R}_{\text{EP}}-\underline{R}_{\text{EP}}}{C_0^{\text{EP}}+C_0^{\text{A}}-\underline{C}_{\text{EP}}-\underline{C}_{\text{A}}} \\
&= \frac{C_0^{\text{EP}}-\underline{C}_{\text{EP}}}{C\left(Q\right)}\times\frac{R_0^{\text{EP}}-\underline{R}_{\text{EP}}}{C_0^{\text{EP}}-\underline{C}_{\text{EP}}}+\frac{C_0^{\text{A}}-\underline{C}_{\text{A}}}{C\left(Q\right)}\times\frac{R_0^{\text{A}}-\underline{R}_{\text{A}}}{C_0^{\text{A}}-\underline{C}_{\text{A}}} \\
&= \frac{C_0^{\text{EP}}-\underline{C}_{\text{EP}}}{C\left(Q\right)}\times\gamma_{\text{Fixe\_HB}}^{\text{EP}}+\frac{C_0^{\text{A}}-\underline{C}_{\text{A}}}{C\left(Q\right)}\times\gamma_{\text{Fixe\_HB}}^{\text{A}}
\end{aligned}
$$

Disaggregation by Groups and Services (EP-G1, EP-G2 and A-G2) :

$$
\begin{aligned}
\frac{c\left(Q_0-\underline{Q}\right)}{C\left(Q\right)}\times\gamma_{\text{Fixe\_HB}} &= \frac{c_{\text{EP}}\left(Q_0-\underline{Q}\right)+c_{\text{A}}\left(Q_0^{\text{A}}-\underline{Q}_{\text{A}}\right)}{C\left(Q\right)}\times\gamma_{\text{Fixe\_HB}} \\
&= \frac{C_0-\underline{C}}{C\left(Q\right)}\times\frac{R_0-\underline{R}}{C_0-\underline{C}} = \frac{C_0-\underline{C}}{C\left(Q\right)}\times\frac{R_{0,1}^{\text{EP}}+R_{0,2}^{\text{EP}}+R_{0,2}^{\text{A}}-\underline{R}_1^{\text{EP}}-\underline{R}_2^{\text{EP}}-\underline{R}_2^{\text{A}}}{C_{0,1}^{\text{EP}}+C_{0,2}^{\text{EP}}+C_{0,2}^{\text{A}}-\underline{C}_{\text{EP}}^1-\underline{C}_{\text{EP}}^2-\underline{C}_{\text{A}}^2} \\
&= \frac{C_{0,1}^{\text{EP}}-\underline{C}_{\text{EP}}^1}{C\left(Q\right)}\times\frac{R_{0,1}^{\text{EP}}-\underline{R}_1^{\text{EP}}}{C_{0,1}^{\text{EP}}-\underline{C}_{\text{EP}}^1}+\frac{C_{0,2}^{\text{EP}}-\underline{C}_{\text{EP}}^2}{C\left(Q\right)}\times\frac{R_{0,2}^{\text{EP}}-\underline{R}_2^{\text{EP}}}{C_{0,2}^{\text{EP}}-\underline{C}_{\text{EP}}^2}+\frac{C_{0,2}^{\text{A}}-\underline{C}_{\text{A}}^2}{C\left(Q\right)}\times\frac{R_{0,2}^{\text{A}}-\underline{R}_2^{\text{A}}}{C_{0,2}^{\text{A}}-\underline{C}_{\text{A}}^2} \\
&= \frac{C_{0,1}^{\text{EP}}-\underline{C}_{\text{EP}}^1}{C\left(Q\right)}\times\gamma_{\text{Fixe\_HB,1}}^{\text{EP}}+\frac{C_{0,2}^{\text{EP}}-\underline{C}_{\text{EP}}^2}{C\left(Q\right)}\times\gamma_{\text{Fixe\_HB,2}}^{\text{EP}}+\frac{C_{0,2}^{\text{A}}-\underline{C}_{\text{A}}^2}{C\left(Q\right)}\times\gamma_{\text{Fixe\_HB,2}}^{\text{A}}
\end{aligned}
$$

(3) As regards the term relating to the variable component the consumption, excluding overconsumption:

$$
\frac{c_{\text{EP}}\left(Q_{\kappa=1}-Q_0\right)+c_{\text{A}}\left(Q_{\kappa=1}^{\text{A}}-Q_0^{\text{A}}\right)}{C\left(Q\right)}\times\gamma_v^{\kappa=1}\equiv T_1
$$

Disaggregation by Groups G1 vs. G2 (customer segments) :

$$
\begin{aligned}
T_1 &= \frac{c_{\text{EP}}\left(Q_{\kappa=1,1}-Q_{0,1}\right)+\left(c_{\text{EP}}+c_{\text{A}}\right)\left(Q_{\kappa=1,2}-Q_{0,2}\right)}{C(Q)}\times\gamma_v^{\kappa=1} \\
&= \frac{c_{\text{EP}}\left(Q_{\kappa=1,1}-Q_{0,1}\right)+c_{\text{EPA}}\left(Q_{\kappa=1,2}-Q_{0,2}\right)}{C(Q)}\times\frac{R_v^{\kappa=1}-R_0}{c_{\text{EP}}\left(Q_{\kappa=1,1}-Q_{0,1}\right)+c_{\text{EPA}}\left(Q_{\kappa=1,2}-Q_{0,2}\right)} \\
&= \frac{c_{\text{EP}}\left(Q_{\kappa=1,1}-Q_{0,1}\right)+c_{\text{EPA}}\left(Q_{\kappa=1,2}-Q_{0,2}\right)}{C(Q)}\times\frac{R_{v,1}^{\kappa=1}+R_{v,2}^{\kappa=1}-R_{0,1}-R_{0,2}}{c_{\text{EP}}\left(Q_{\kappa=1,1}-Q_{0,1}\right)+c_{\text{EPA}}\left(Q_{\kappa=1,2}-Q_{0,2}\right)} \\
&= \frac{c_{\text{EP}}\left(Q_{\kappa=1,1}-Q_{0,1}\right)}{C(Q)}\times\frac{R_{v,1}^{\kappa=1}-R_{0,1}}{c_{\text{EP}}\left(Q_{\kappa=1,1}-Q_{0,1}\right)}+\frac{c_{\text{EPA}}\left(Q_{\kappa=1,2}-Q_{0,2}\right)}{C(Q)}\times\frac{R_{v,2}^{\kappa=1}-R_{0,2}}{c_{\text{EPA}}\left(Q_{\kappa=1,2}-Q_{0,2}\right)} \\
&= \frac{c_{\text{EP}}\left(Q_{\kappa=1,1}-Q_{0,1}\right)}{C(Q)}\times\gamma_{v,1}^{\kappa=1}+\frac{c_{\text{EPA}}\left(Q_{\kappa=1,2}-Q_{0,2}\right)}{C(Q)}\times\gamma_{v,2}^{\kappa=1}
\end{aligned}
$$

Disaggregation by Services (EP vs. A) :

$$
\begin{aligned}
T_1 &== \frac{c_{\text{EP}}\left(Q_{\kappa=1}-Q_0\right)+c_{\text{A}}\left(Q_{\kappa=1}^{\text{A}}-Q_0^{\text{A}}\right)}{C(Q)}\times\frac{R_v^{\kappa=1}-R_0}{c_{\text{EP}}\left(Q_{\kappa=1}-Q_0\right)+c_{\text{A}}\left(Q_{\kappa=1}^{\text{A}}-Q_0^{\text{A}}\right)} \\
&= \frac{c_{\text{EP}}\left(Q_{\kappa=1}-Q_0\right)+c_{\text{A}}\left(Q_{\kappa=1}^{\text{A}}-Q_0^{\text{A}}\right)}{C(Q)}\times\frac{R_{v,\text{EP}}^{\kappa=1}+R_{v,\text{A}}^{\kappa=1}-R_{0,\text{EP}}-R_{0,\text{A}}}{c_{\text{EP}}\left(Q_{\kappa=1,1}+Q_{\kappa=1,2}-Q_{0,1}-Q_{0,2}\right)+c_{\text{A}}\left(Q_{\kappa=1,2}^{\text{A}}-Q_{0,2}^{\text{A}}\right)} \\
&= \frac{c_{\text{EP}}\left(Q_{\kappa=1}-Q_0\right)}{C(Q)}\times\frac{R_{v,\text{EP}}^{\kappa=1}-R_{0,\text{EP}}}{c_{\text{EP}}\left(Q_{\kappa=1}-Q_0\right)}+\frac{c_{\text{A}}\left(Q_{\kappa=1}^{\text{A}}-Q_0^{\text{A}}\right)}{C(Q)}\times\frac{R_{v,\text{A}}^{\kappa=1}-R_{0,\text{A}}}{c_{\text{A}}\left(Q_{\kappa=1,2}-Q_{0,2}\right)} \\
&= \frac{c_{\text{EP}}\left(Q_{\kappa=1}-Q_0\right)}{C(Q)}\times\gamma_{v,\text{EP}}^{\kappa=1}+\frac{c_{\text{A}}\left(Q_{\kappa=1}^{\text{A}}-Q_0^{\text{A}}\right)}{C(Q)}\times\gamma_{v,\text{A}}^{\kappa=1}
\end{aligned}
$$

Disaggregation by Groups and Services (EP-G1, EP-G2 and A-G2) :

$$
\begin{aligned}
T_1 &= \frac{c_{\text{EP}}\left(Q_{\kappa=1}-Q_0\right)+c_{\text{A}}\left(Q_{\kappa=1}^{\text{A}}-Q_0^{\text{A}}\right)}{C(Q)}\times\frac{R_v^{\kappa=1}-R_0}{c_{\text{EP}}\left(Q_{\kappa=1}-Q_0\right)+c_{\text{A}}\left(Q_{\kappa=1}^{\text{A}}-Q_0^{\text{A}}\right)} \\
&= \frac{c_{\text{EP}}\left(Q_{\kappa=1}-Q_0\right)+c_{\text{A}}\left(Q_{\kappa=1}^{\text{A}}-Q_0^{\text{A}}\right)}{C(Q)}\times\frac{R_{v,1}^{\kappa=1,\text{EP}}+R_{v,2}^{\kappa=1,\text{EP}}+R_{v,2}^{\kappa=1,\text{A}}-R_{0,\text{EP}}^{1}-R_{0,\text{EP}}^{2}-R_{0,\text{A}}^{2}}{c_{\text{EP}}\left(Q_{\kappa=1,1}+Q_{\kappa=1,2}-Q_{0,1}-Q_{0,2}\right)+c_{\text{A}}\left(Q_{\kappa=1,2}^{\text{A}}-Q_{0,2}^{\text{A}}\right)} \\
&= \frac{c_{\text{EP}}\left(Q_{\kappa=1,1}-Q_{0,1}\right)}{C(Q)}\times\frac{R_{v,1}^{\kappa=1,\text{EP}}-R_{0,\text{EP}}^{1}}{c_{\text{EP}}\left(Q_{\kappa=1,1}-Q_{0,1}\right)}+\frac{c_{\text{EP}}\left(Q_{\kappa=1,2}-Q_{0,2}\right)}{C(Q)}\times\frac{R_{v,2}^{\kappa=1,\text{EP}}-R_{0,\text{EP}}^{2}}{c_{\text{EP}}\left(Q_{\kappa=1,2}-Q_{0,2}\right)} \\
&\quad+\frac{c_{\text{A}}\left(Q_{\kappa=1,2}^{\text{A}}-Q_{0,2}^{\text{A}}\right)}{C(Q)}\times\frac{R_{v,2}^{\kappa=1,\text{A}}-R_{0,\text{A}}^{2}}{c_{\text{A}}\left(Q_{\kappa=1,2}^{\text{A}}-Q_{0,2}^{\text{A}}\right)} \\
&= \frac{c_{\text{EP}}\left(Q_{\kappa=1,1}-Q_{0,1}\right)}{C(Q)}\times\gamma_{v,\text{EP},1}^{\kappa=1}+\frac{c_{\text{EP}}\left(Q_{\kappa=1,2}-Q_{0,2}\right)}{C(Q)}\times\gamma_{v,\text{EP},1}^{\kappa=1}+\frac{c_{\text{A}}\left(Q_{\kappa=1,2}^{\text{A}}-Q_{0,2}^{\text{A}}\right)}{C(Q)}\times\gamma_{v,\text{A},2}^{\kappa=1}
\end{aligned}
$$

(4) As regards the term relating to overconsumption (linked to tariff misperception):

$$T_2 \equiv \frac{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0} - Q_{\kappa=1}\right) + c_{\mathrm{A}}\left(Q^{\mathrm{A}}_{\kappa=\kappa_0} - Q^{\mathrm{A}}_{\kappa=1}\right)}{C(Q)} \times \gamma_{\mathrm{surco}}$$

Disaggregation by Groups G1 vs. G2 (customer segments) :

$$T_2 = \frac{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0,1} - Q_{\kappa=1,1}\right) + \left(c_{\mathrm{EP}} + c_{\mathrm{A}}\right)\left(Q_{\kappa=\kappa_0,2} - Q_{\kappa=1,2}\right)}{C(Q)} \times \frac{R_v^{\kappa=\kappa_0} - R_v^{\kappa=1}}{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0,1} - Q_{\kappa=1,1}\right) + c_{\mathrm{EPA}}\left(Q_{\kappa=\kappa_0,2} - Q_{\kappa=1,2}\right)}$$

$$= \frac{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0,1} - Q_{\kappa=1,1}\right) + c_{\mathrm{EPA}}\left(Q_{\kappa=\kappa_0,2} - Q_{\kappa=1,2}\right)}{C(Q)} \times \frac{R_{v,1}^{\kappa=\kappa_0} + R_{v,2}^{\kappa=\kappa_0} - R_{v,1}^{\kappa=1} - R_{v,2}^{\kappa=1}}{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0,1} - Q_{\kappa=1,1}\right) + c_{\mathrm{EPA}}\left(Q_{\kappa=\kappa_0,2} - Q_{\kappa=1,2}\right)}$$

$$= \frac{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0,1} - Q_{\kappa=1,1}\right)}{C(Q)} \times \frac{R_{v,1}^{\kappa=\kappa_0} - R_{v,1}^{\kappa=1}}{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0,1} - Q_{\kappa=1,1}\right)} + \frac{c_{\mathrm{EPA}}\left(Q_{\kappa=\kappa_0,2} - Q_{\kappa=1,2}\right)}{C(Q)} \times \frac{R_{v,2}^{\kappa=\kappa_0} - R_{v,2}^{\kappa=1}}{c_{\mathrm{EPA}}\left(Q_{\kappa=\kappa_0,2} - Q_{\kappa=1,2}\right)}$$

$$= \frac{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0,1} - Q_{\kappa=1,1}\right)}{C(Q)} \times \gamma_{\mathrm{surco},1} + \frac{c_{\mathrm{EPA}}\left(Q_{\kappa=\kappa_0,2} - Q_{\kappa=1,2}\right)}{C(Q)} \times \gamma_{\mathrm{surco},2}$$

Disaggregation by Services (EP vs. A) :

$$T_2 = \frac{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0} - Q_{\kappa=1}\right) + c_{\mathrm{A}}\left(Q^{\mathrm{A}}_{\kappa=\kappa_0} - Q^{\mathrm{A}}_{\kappa=1}\right)}{C(Q)} \times \frac{R_v^{\kappa=\kappa_0} - R_v^{\kappa=1}}{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0} - Q_{\kappa=1}\right) + c_{\mathrm{A}}\left(Q^{\mathrm{A}}_{\kappa=\kappa_0} - Q^{\mathrm{A}}_{\kappa=1}\right)}$$

$$= \frac{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0} - Q_{\kappa=1}\right) + c_{\mathrm{A}}\left(Q^{\mathrm{A}}_{\kappa=\kappa_0} - Q^{\mathrm{A}}_{\kappa=1}\right)}{C(Q)} \times \frac{R_{v,\mathrm{EP}}^{\kappa=\kappa_0} + R_{v,\mathrm{A}}^{\kappa=\kappa_0} - R_{v,\mathrm{EP}}^{\kappa=1} - R_{v,\mathrm{A}}^{\kappa=1}}{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0} - Q_{\kappa=1}\right) + c_{\mathrm{A}}\left(Q^{\mathrm{A}}_{\kappa=\kappa_0} - Q^{\mathrm{A}}_{\kappa=1}\right)}$$

$$= \frac{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0} - Q_{\kappa=1}\right)}{C(Q)} \times \frac{R_{v,\mathrm{EP}}^{\kappa=\kappa_0} - R_{v,\mathrm{EP}}^{\kappa=1}}{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0} - Q_{\kappa=1}\right)} + \frac{c_{\mathrm{A}}\left(Q^{\mathrm{A}}_{\kappa=\kappa_0} - Q^{\mathrm{A}}_{\kappa=1}\right)}{C(Q)} \times \frac{R_{v,\mathrm{A}}^{\kappa=\kappa_0} - R_{v,\mathrm{A}}^{\kappa=1}}{c_{\mathrm{A}}\left(Q^{\mathrm{A}}_{\kappa=\kappa_0} - Q^{\mathrm{A}}_{\kappa=1}\right)}$$

$$= \frac{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0} - Q_{\kappa=1}\right)}{C(Q)} \times \gamma^{\mathrm{EP}}_{\mathrm{surco}} + \frac{c_{\mathrm{A}}\left(Q^{\mathrm{A}}_{\kappa=\kappa_0} - Q^{\mathrm{A}}_{\kappa=1}\right)}{C(Q)} \times \gamma^{\mathrm{A}}_{\mathrm{surco}}$$

Disaggregation by Groups and Services (EP-G1, EP-G2 and A-G2) :

$$T_2 = \frac{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0,1} - Q_{\kappa=1,1}\right) + c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0,2} - Q_{\kappa=1,2}\right) + c_{\mathrm{A}}\left(Q^{\mathrm{A}}_{\kappa=\kappa_0,2} - Q^{\mathrm{A}}_{\kappa=1,2}\right)}{C(Q)} \times \gamma_{\mathrm{surco}}$$

with:

$$\gamma_{\mathrm{surco}} = \frac{R_v^{\kappa=\kappa_0} - R_v^{\kappa=1}}{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0,1} - Q_{\kappa=1,1}\right) + c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0,2} - Q_{\kappa=1,2}\right) + c_{\mathrm{A}}\left(Q^{\mathrm{A}}_{\kappa=\kappa_0,2} - Q^{\mathrm{A}}_{\kappa=1,2}\right)}$$

$$= \frac{R_{v,1}^{\kappa=\kappa_0,\mathrm{EP}} + R_{v,2}^{\kappa=\kappa_0,\mathrm{EP}} + R_{v,2}^{\kappa=\kappa_0,\mathrm{A}} - R_{v,1}^{\kappa=1,\mathrm{EP}} - R_{v,2}^{\kappa=1,\mathrm{EP}} - R_{v,2}^{\kappa=1,\mathrm{A}}}{c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0,1} - Q_{\kappa=1,1}\right) + c_{\mathrm{EP}}\left(Q_{\kappa=\kappa_0,2} - Q_{\kappa=1,2}\right) + c_{\mathrm{A}}\left(Q^{\mathrm{A}}_{\kappa=\kappa_0,2} - Q^{\mathrm{A}}_{\kappa=1,2}\right)}$$

Inserting this expression into the previous one and rearranging, we get:

$$T_2 \equiv \frac{c_{\text{EP}}\left(Q_{\kappa=\kappa_0,1} - Q_{\kappa=1,1}\right)}{C(Q)} \times \gamma^{\text{EP}}_{\text{surco},1} + \frac{c_{\text{EP}}\left(Q_{\kappa=\kappa_0,2} - Q_{\kappa=1,2}\right)}{C(Q)} \times \gamma^{\text{EP}}_{\text{surco},2} + \frac{c_{\text{A}}\left(Q^{\text{A}}_{\kappa=\kappa_0,2} - Q^{\text{A}}_{\kappa=1,2}\right)}{C(Q)} \times \gamma^{\text{A}}_{\text{surco},2}$$

with :

$$\gamma^{\text{EP}}_{\text{surco},1} = \frac{R^{\text{EP},\kappa=\kappa_0}_{v,1} - R^{\text{EP},\kappa=1}_{v,1}}{c_{\text{EP}}\left(Q_{\kappa=\kappa_0,1} - Q_{\kappa=1,1}\right)}$$

$$\gamma^{\text{EP}}_{\text{surco},2} = \frac{R^{\text{EP},\kappa=\kappa_0}_{v,2} - R^{\text{EP},\kappa=1}_{v,2}}{c_{\text{EP}}\left(Q_{\kappa=\kappa_0,2} - Q_{\kappa=1,2}\right)}$$

$$\gamma^{\text{A}}_{\text{surco},2} = \frac{R^{\text{A},\kappa=\kappa_0}_{v,1} - R^{\text{A},\kappa=1}_{v,1}}{c_{\text{A}}\left(Q_{\kappa=\kappa_0,2} - Q_{\kappa=1,2}\right)}$$

■

MAIN VARIABLES (LIST)

A) Consumption

q_i : household water consumption *i*

q_{ij} : water consumption of household *i* in block *j*

q_i^{IBT} : water consumption of household *i* facing the IBT tariff system

$q_i^{\text{IBT-PP}}$: water consumption of household *i* facing the IBT system when the IBT system is properly perceived

q_i^{TBSE} : water consumption of household *i* facing the TBSE system

$q_i^d(\cdot)$: demand function (for drinking water utility or drinking water and wastewater utility) of household *i*

$q_{ij}^d(\cdot)$: conditional demand function of household *i* in block *j*

q_{0i} : captive consumption of household *i*

q_{0i}^j : captive consumption of household *i* in block *j*

$\underline{q}_i$: basic consumption of household *i*

$\underline{q}_i^j$: basic consumption of household *i* in block *j*

$q_i - q_{0,i} = (q - q_0)_i$: variable (or "economic") part of of household i's water consumption

$(q - q_0)_i^j$: variable (or "economic") part of of household i's water consumption in block *j*

$q_i - \underline{q}_i$: non-basic water consumption of household *i*

$(q - \underline{q})_i^j$ non-basic water consumption of household *i* in block *j*

$\underline{\text{B}}_i$, B_i^0, B_i^{IBT} et $\text{B}_i^{\text{IBT-PP}}$: discriete variable indicating the number of the consumption block in which basic consumption, captive consumption, IBT consumption and IBT consumption with perfect perception are located

Q : total consumption (and production, referred as "service level")

Q_{EP} : total consumption for the drinking water service

Q_{A} : total consommation for collective wastewater treatment service

B) Cost (Operator)

CF_{EP} : level of fixed costs (borne by the operator) for the drinking water service

$\mathrm{CVM}_{\mathrm{EP}} = c_{\mathrm{EP}}$: unit variable cost (assumed to be constant) for the supply of one cubic metre of drinking water (also equal to the marginal cost of drinking water service)

CF_{A} : level of fixed costs (borne by the operator) for the collective wastewater service

$\mathrm{CVM}_{\mathrm{A}} = c_{\mathrm{A}}$: unit variable cost (assumed to be constant) for the treatment of one cubic metre of domestic wastewater (also equal to the marginal cost of wasterwater treatment service)

$c_{\mathrm{EPA}} = c_{\mathrm{EP}} + c_{\mathrm{A}}$ the unit variable cost for the "EPA" service (production of one clean domestic water cubic meter)

c_e : value of the environmental cost (in euros per cubic metre), defined as the cost of fully depolluting one cubic metre of domestic waste water

n : Number of domestic customers (households)

$n = n_{\mathrm{EP}}$: number of domestic subscribers (households) to the public drinking water service

n_{A} : number of domestic subscribers (households) to the public sewerage service with $n_{\mathrm{A}} \leq n = n_{\mathrm{EP}}$

$n_1 = n - n_{\mathrm{A}}$: Size of the G1 group of domestic customers (households) who are not connected to the public sewerage system / only pay for drinking water service

$n_2 = n_{\mathrm{A}}$: Size of the G2 group of domestic customers (households) who are connected to the public sewerage system / pays for the drinking water and wastewater service

$C_{\mathrm{EP}} = C_{\mathrm{EP}}(Q_{\mathrm{EP}})$: Operator cost function for the drinking water service

$C_{\mathrm{A}} = C_{\mathrm{A}}(Q_{\mathrm{A}})$: Operator cost function for the wastewater service

$C(Q) = C_{\mathrm{EP}}(Q_{\mathrm{EP}}) + C_{\mathrm{A}}(Q_{\mathrm{A}})$: Operator cost function for the general (drinking water / drinking water and wastewater) service

$C_i^{\mathrm{EP}}(q_i)$: cost to the service (borne by the operator) of providing the drinking water service to household *i* whose consumption is q_i

$C_i^{\mathrm{A}}(q_i)$: cost to the service (borne by the operator) of providing the collective sanitation service to household *i* whose consumption is q_i

$C_0 = CF + cQ_0$: cost to the service of satisfying captive consumption (fixed costs in the economic sense of the term)

$\underline{C} = CF + c\underline{Q}$: cost to service for the provision of the total basic service

γ , γ_0 , γ_v … coverage rate of the cost of the service (general, for the satisfaction of captive consumption, for the satisfaction of the variable part of consumption ...) by the marketing of the corresponding units of service

C) Tariff

$T_i = T(q_i^d(\cdot))$: Tariff function

$R_0 = \sum_i T(q_{0i})$: revenue from the marketing of captive consumption Q_0

$\underline{R} = \sum_i T(\underline{q}_i)$: revenue from the marketing of basic service (includes the collection of fixed part (subscriptions)).

F : fixed part (amount of the subscription); F_{EP} : subscription amount for the drinking water service, F_{A} : subscription amount for the collective wastewater service

F_{TBSE} : subscription amount in the TBSE system ($F_{\mathrm{EP}}^{\mathrm{TBSE}}$: subscription amount for the drinking water supply in the TBSE system, $F_{\mathrm{A}}^{\mathrm{TBSE}}$: subscription amount for the wastewater supply in the TBSE system)

π : price per cubic metre (Two-part Tariff), π_{EP} : price per cubic metre for the drinking water supply (Two-part Tariff), π_{A} : price per cubic metre for the collective wastewater supply (Two-part Tariff).

π_{TBSE} : price per cubic metre in the TBSE system (Two-part Tariff), $\pi_{\mathrm{EP}}^{\mathrm{TBSE}}$: price per cubic metre in the TBSE system for the drinking water supply, $\pi_{\mathrm{A}}^{\mathrm{TBSE}}$: price per cubic metre in the TBSE system for the collective wastewater supply

p : number of consumption blocks (IBT)

$I_1 = [0, k_1]$, $I_2 =]k_1, k_2]$, $I_3 =]k_2, k_3]$ … : consumption blocks (IBT)

k_1, k_2, k_3 … : tariff thresholds (IBT)

π_1, π_2, π_3 … : unit prices (per m^3) in consumption block $I_1 = [0, k_1]$,, $I_2 =]k_1, k_2]$, $I_3 =]k_2, k_3]$ … with $\pi_1 < \pi_2 < \pi_3 < \ldots$ (IBT)

(F, π) vector of tariff parameters for Two Part tariff

$\theta_2 = \left(F, \pi_1, \pi_2, k_1\right)$ vector of tariff parameters for an IBT2, $\theta_3 = \left(F, \pi_1, \pi_2, \pi_3, k_1, k_2\right)$ vector of tariff parameters for an IBT3, …

$\pi = \pi\left(q\right)$: marginal price function (unit price scale)

$D = D\left(q\right)$: Nordin' D function (also called "Difference variable" function)

$D_1 = 0$: Nordin's D value of block 1, $D_2 = \left(\pi_2 - \pi_1\right)\cdot k_1$ **:** Nordin's D value of block 2, $D_3 = D_2 + \left(\pi_3 - \pi_2\right)\cdot k_2$ **:** Nordin's D value of block 3 …

$\overline{\pi}$: average price of consumption, calculated excluding subscription fees

$\overline{k}$: tariff threshold (structural progressivity)

D) Invoices

t : VAT rate (t_{EP} : VAT rate for the drinking water service, t_{A} : VAT rate for the wastewater service)

r : amount of the excise duty (in euros per cubic metre), r_{EP} : amount of the excise duty for the drinking water service, r_{A} : amount of the excise duty for the collective wastewater service

T_i^{IBT}, $T_i^{\text{IBT-PP}}$, T_i^{TBSE} : amount of the water bill in relation to the IBT system, the IBT system when it is perfectly perceived, the TBSE system

$T_i^{\text{IBT}} - T_i^{\text{IBT-PP}}$: IBT mismanagement cost (linked to tarif misperception)

$T_{0i} = T(q_{0i})$; captive part of the water bill ($T_{i,0}^{\text{IBT}}$, $T_{i,0}^{\text{TBSE}}$)

$T_i - T_{0i}$: non-captive part of the water bill ($T_i^{\text{IBT}} - T_{i,0}^{\text{IBT}}$,, $T_i^{\text{IBT-PP}} - T_{i,0}^{\text{IBT}}$, $T_i^{\text{TBSE}} - T_{i,0}^{\text{TBSE}}$)

$\underline{T}_i = T(\underline{q}_i)$: minimum basic part of the water bill ($\underline{T}_i^{\text{IBT}} = T_{\text{IBT}}(\underline{q}_i)$, $\underline{T}_i^{\text{TBSE}} = T_{\text{TBSE}}(\underline{q}_i)$)

$T_i - \underline{T}_i$: non basic part of the water bill (($T_i^{\text{IBT}} - \underline{T}_i^{\text{IBT}}$, $T_i^{\text{IBT-PP}} - \underline{T}_i^{\text{IBT}}$, $T_i^{\text{TBSE}} - \underline{T}_i^{\text{TBSE}}$)

$T_{i,1}^{\text{IBT}}$, $T_{i,2}^{\text{IBT}}$ … : IBT expenditure for consumption units in Block 1, in Block 2 …

$T_{i,1}^{\text{IBT-PP}}$, $T_{i,2}^{\text{IBT-PP}}$ … IBT expenditure for consumption units in Block 1, in Block 2 … when the IBT is perfectly perceived

$\underline{T}_{i,1}^{\text{IBT}}$, $\underline{T}_{i,2}^{\text{IBT}}$ … IBT expenditure for basic consumption units in Block 1, in Block 2 …

$\overline{T}\left(q\right)$: average cost of consumption (Household)

$T'\left(q\right)$: marginal cost of consumption (Household)

E) Subventions – "Taxations" (Marges)

$\sigma_j = (\pi_j - c)_- = -\min[\pi_j - c, 0]$: subsidy rate (in euros per cubic metre) in consumption block *j*

$\tau_j = (\pi_j - c)_+ = \max[\pi_j - c, 0]$: margin per unit of service (in euros per cubic metre) in consumption block *j*

c_{i0} : subsidy/taxation of the access fee, c_{i0}^+ tax amount (margin) on the Acess Fee, c_{i0}^- subsidy amount on Access Fee

$c_{ij} = (\pi_j - c) \times q_i^j$: subsidy/taxation on household *i* in consumption in block *j*, c_{ij}^+ : amount of tax (margin) levied on household *i*'s consumption in block *j*, c_{ij}^- : amount of subsidy received by household *i* on consumption of tranche *j*

s_{iq} : amount of subsidy received by household *i* on its consumption q_i

s_i : total subsidy received by the household *i*, including Access Fee

t_{iq} : amount of 'taxation' (margin) levied on the consumption of household *i*

t_i : total tax (margin) levied on household *i*, including Access Fee

S_q : Total (mass) of subsidies granted on total household water consumption

S : Total (mass) of subsidies granted on households, including Access Fee

T_q : Total (mass) of taxes (margins) levied on total household water consumption

T : Total (mass) of taxes (margins) levied on the domestic customer portfolio

$\underline{c}_{ij} = (\pi_j - c) \times \underline{q}_i^j$ subsidy/taxation on household *i*'s basic consumption in block *j*, $\underline{c}_{ij}^+$: amount of the taxation (margin) levied on the basic consumption of household *i* in block *j*, $\underline{c}_{ij}^-$: amount of the subsidy paid on the basic consumption of household *i* in block *j*,

$\underline{s}_{iq}$: amount of subsidy received by household *i* on its basic consumption $\underline{q}_i$

$\underline{s}_i$: total subsidy received by household *i* for the basic service (including access fee)

$\underline{t}_{iq}$: amount of "taxation" (margin) generated on the basic consumption $\underline{q}_i$ of household *i*

$\underline{t}_i$: total margin generated on the provision of the basic service to household *i*, including Access Fee

$S_{\underline{q}}$: Total (mass) of subsidies granted on total basic household consumption

$\underline{S}$: Total (mass) of subsidies paid to households for the provision of basic service, including Access Fee

$T_{\underline{q}}$: Total (mass) of margins generated on total basic household consumption

$\underline{T}$: Total (mass) of margins generated on the provision of basic service to households, including Access Fee

…

Π : profit (operating result)

Π_i : profit (operating result) on Household *i*

$\Pi_{0i} = F - \frac{CF}{n}$: margin on Access Fee

m_i : margin on consumption q_i of Household i

$\hat{\Pi}_{i0} = F - \frac{CF}{n} - D_i$: : pseudo-subsidy / taxation (margin) on the Access Fee

$\hat{m}_i = (\pi_i - c) \times q_i$: pseudo-subsidy / taxation (margin) on consumption (with $\pi_i = \pi(q_i)$ the value of the marginal price faced by household *i*)

$M_0 = R_0 - C_0$: total net margin (potentially negative) generated on total captive consumption Q_0,

$\underline{M} = \underline{R} - \underline{C}$: total net margin (potentially negative) generated on the provision of the total basic service

$\mathrm{CF} - nF \quad \frac{\mathrm{R}}{Q} \quad \mathrm{RVM} \quad \mathrm{R} \quad R \quad \frac{R}{Q} \quad \frac{C}{Q} \quad \frac{\mathrm{CF}}{n} - F \quad \frac{m_q}{\overline{q}} \quad n \times m_q \quad \Pi \quad \frac{\Pi}{Q}$

Q_+ : total subsidised consumption

Q_- : total consumption that is not subsidised (possibly margined)

$\underline{Q}^+$: total subsidised basic consumption

$\underline{Q}^-$: total unsubsidised basic consumption (possibly margined)

$(Q - \underline{Q})^+$: total subsidised non-basic consumption

$(Q - \underline{Q})^-$: total unsubsidised non-basic consumption (possibly margined)

$\underline{q}_i^+$: subsidised basic consumption of household *i*

$(q_i - \underline{q}_i)^+$: subsidised non-basic consumption of household *i*,

$\underline{q}_i^-$: basic consumption of household *i* that is not subsidised,

$\left(q_i - \underline{q}_i\right)^-$: non-basic consumption of household *i* that is not subsidised,

F) Statistics (x denotes the variable of interest)

$\bar{x}$: average (mean)

$\tilde{x}$: median

D1, D2, …, D9 : first decile, second decile, ... last decile

Q1, Q3 : first quartile, third quartile

$F = F(x)$: distribution function

$V(x) = \sigma_x^2$: variance

σ_x : standard deviation

MAPE : average of deviations from the arithmetic mean

$i(x)$: Gini index

$S = S(x)$: Schutz coefficient

Ω ratio : ratio of the average of the variable of interest calculated for the sole population of 'poor' units (whose standard of living is below the poverty line) to the average of the variable of interest calculated for the population as a whole

PPV (positive predictive value) : share of basic consumption in subsidised units

FDR : share of non-basic consumption in subsidised units

FOR : share of basic consumption in units that are not subsidised (possibly margined)

NPV : share of non basic consumption in units that are not subsidised (possibly margined)

p (prevalence rate) : proportion of production Q aiming to meet total basic needs, $\underline{Q}$

LR_+ : odds within subsidised units (for a non basic unit wrongly subsidised, how many basic units are rightly subsidised)

LR_- : odds within unsubsidised units (for a non basic unit righly unsubsidised, how many basic units are wrongly unsubsidised)

DOR (Diagnostic Odd Ratio) : odds ratio

TPR (True Positive Rate) : proportion of basic consumption that is subsidised

FPR (False Positive Rate) : proportion of non-basic consusmmation that is subsidised

FNR (False Negative Rate) : proportion of basic consumption that is not subsidised (possibly margined)

TNR (True Négative Rate) : proportion of non basic consumption that is not subsidised (possibly margined)

G) Demographic and socio-economic variables

R_i : net income of household *i*

$\eta_{y,x}$: elasticity of the variable y (dependent variable, explained variable) in relation to the variable x (explanatory variable, determinant)

κ : tariff perception parameter

CAR_i : weight of EP/EPA bill in household income for household *i*

PAR_i : weight of EP/EPA bill in household income *i* to cover basic needs

e_i : affordability deficit of household *i* (e_i^{PAR} : affordability deficit as defined by the PAR for household *i*, e_i^{CAR} : affordability deficit as defined by the CAR for household *i*)

$H_{\mathrm{PAR}}^{\mathrm{Household}}$: PAR Household Headcount ratio (proportion of households facing an affordability issue as defined by the PAR criterion)

$H_{\mathrm{CAR}}^{\mathrm{Household}}$: CAR Household Headcount ratio (proportion of households facing an affordability issue as defined by the CAR criterion)

Γ : aggregate surplus (Welfare)

$U_i(\cdot,\cdot)$: utility function of household *i*

$u_i(\cdot)$: Gross surplus function of household *i*

$v_i(\cdot)$: net surplus function (also known as consumer surplus) of household *i* (with then $v_i(\cdot) = u_i(\cdot) - T(\cdot)$)

γ_i : contribution to aggregate surplus Γ of household *i*

N_i : family size of Household *i*

SNWA_i : share of non-working adults with respect to total number of adults within Household *i*

Pool : dummy variable which takes the value 1 if the household has a swimming pool (0 otherwise),

Garden : dummy variable which takes the value 1 if the household has a garden (0 otherwise),

Weather : percentage of days without rain over the billing period (rainfall frequency).

REFERENCES

Alivon, F. [2016], "La ségrégation spatiale et économique Une analyse en termes d'emploi et d'éducation dans les espaces urbains", Doctoral dissertation, Université de Bourgogne

Almendarez-Hernández M.A., Polanco G.A., Trejo V.H., Ortega-Rubio A. & Morales L.F.B. [2016], "Residential Water Demand in a Mexican Biosphere Reserve: Evidence of the Effects of Perceived Price", Water 8 (10), pp. 1-14.

Arbuès F., Garcia-Valiñas M.A. & Martinez-Espiñeira R. [2003], "Estimation of residential water demand: a state-of-art review", Journal of Socio-Economics, 32, pp. 81-102.

ARS Réunion [2022], Bilan 2021, Qualité de l'eau du robinet à La Réunion.

Atkinson A.B. [1970], "On the measurement of inequality", Journal of Economic Theory, 2 (3), pp. 244-263.

Barberán R. & Arbués F. [2008], "Equity in domestic water rates design", Water Resources Management, 23, pp. 2101-2118.

Barberán R., Arbués F. & Domínguez F. [2006], Consumo y Gravamen del Agua para Usos Residenciales en la Ciudad de Zaragoza. Evaluación y Propuesta de Reforma, Zaragoza: Ayuntamiento de Zaragoza, Servicio de Cultura.

Barde J.A. & Lehmann P. [2014], "Distributional effects of water tariff reforms - An empirical study for Lima, Peru", Water Resources and Economics, 6, pp. 30-57.

Bell W. [1954], "A probability model of the measurement of ecological segregation", Social Forces, 32(4), pp. 357-364.

Bhargave S., Loewenstein G. & Sydnor J. [2017], "Choose to Lose: Health Plan Choices from a Menu with Dominated Options", The Quarterly Journal of Economics, 132, pp. 1319-1372.

Binet M.E., Carlevaro F., & Paul M. [2016], La demande d'eau potable des ménages à La Réunion : que nous apprennent les études économétriques, in Droit, Economie et Gestion de l'Eau dans la Zone Océan Indien, pp. 30-51.

Binet M.E., Carlevaro F. & Paul M. [2014], "Estimation of Residential Water Demand with Imperfect Price Perception", Environmental and Resource Economics, 59 (4), pp. 561-581.

Brick K., De Martino S. & Visser M. [2017], "Behavioural Nudges for Water Conservation: Experimental Evidence from Cape Town"[73].

Cartter D.W. & Milon J. [2005], "Price Knowledge in Household Demand for Utility Services", Land Economics, 81, pp. 265-283.

Carlevaro F., Schlesser C., Binet M.E., Durand S. & Paul M. [2007], "Econometric modeling and analysis of residential water demand based on unbalanced panel data", Pricladniaya Ekonometrica (Applied Econometrics), Market DS, Moscow 4 (8), pp. 81-100.

[73] https://www.researchgate.net/publication

Cavanagh S. M., Hanemann M. & Stavins R. N. [2002], "Muffled Price Signals: Household Water Demand Under Increasing-Block Prices"[74].

Chetty R. & Saez E. [2013], "Teaching the Tax Code: Earnings Responses to an Experiment with EITC Recipients", American Economic Journal: Applied Economics, 5, pp. 1-31.

Cohen J., "A coefficient of agreement for nominal scales", Educ. Psychol. Meas. 1960, 20, pp. 27-46.

Crampes C. & Lozachmeur J.M. [2014], "Progressive tariff, efficiency and equity", Revue d'économie industrielle, 148, 4th quarter 2014, Energy transition, industries and markets.

Dagum C. [1997], "A New Approach to the Decomposition of the Gini Income Inequality Ratio", Empirical Economics, 22 (4), pp. 515-531.

Dahan M. & Nisan U. [2013], "The Unintended Consequences of IIBT Pricing Policy in Urban Water", Water Resources Research 43, W03402.

Dalhuisen, J. M., Florax, R., de Groot, H. & Nijkamp P. [2003], "Price and income elasticities of residential water demand: A meta-analysis", Land Economics 79 (2), pp. 292-308.

Duncan O.D. et Duncan B. [1955a], "A methodological analysis of segregation indexes", American Sociological Review, 20(2), pp. 210-217.

Duncan O.D. et Duncan B. [1955b], "Residential distribution and occupational stratification", American Journal of Sociology,60, pp. 493-503.

Duncan O.D., Cuzzort R.P. et Duncan B. [1961], Statistical geography: Problems in analyzing areal data, The Free Press of Glencoe, Illinois.

Estupiñán N., Gómez-Lobo A., Muñoz-Raskin R. & Serebrisky T. [2007], "Affordability and Subsidies in Public Urban Transport: What Do We Mean, What Can Be Done?", World Bank.

Fankhauser S. & Tepic S. [2007], "Can poor consumers pay for energy and water? An affordability analysis for transition countries", Energy Policy 35, pp. 1038-1049.

Foster V. & Yepes T. [2006], "Is Cost Recovery a Feasible Objective for Water and Electricity? The Latin American Experience", Policy Research Working Paper 3943, World Bank,Washington, DC.

Fuente D, Gakii Gatua J., Ikiara M., Kabubo-Mariara J., Mwaura M. & Whittington D. [2016], "Water and sanitation service delivery, pricing, and the poor: An empirical estimate of subsidy incidence in Nairobi, Kenya", Water Resources Research, vol. 52, Issue 6, pp. 4845-4862.

García-Valiñas M, Martínez-Espiñeira R & González-Gómez F [2010], "Affordability of residential water tariffs: Alternative measurement and explanatory factors in southern Spain", Journal of Environmental Management, vol 91, pp. 2696-2706.

García-Valiñas M, Nauges C & Reynaud A [2009], "How much water do residential users really need? An estimation of minimum water requirements for French households", unpublished manuscript.

[74] Available at SSRN: https://ssrn.com/abstract=317924

Garcia-Valiñas M., Athukorala W., Wilson C., Torgler, B. & Gifford R. [2014], "Nondiscretionary residential water use: The impact of habits and water-efficient technologies". Australian Journal of Agricultural and Resource Economics, 58, 185-204.

Gaudin S. [2006], "Effect of Price Information on Residential Water Demand", Applied Economics, 38 (4), pp. 383-393.

Gaudin S., Griffin R.C. & Sickles R.C. [2001], "Demand specification for municipal water management: Evaluation of the Stone-Geary form", Land Economics, 77 (3), pp. 399-422.

Gómez-Lobo A. & Contreras D. [2003], "Water Subsidy Policies: a comparison of the Chilean and Colombian schemes", World Bank Economic Review, vol. 17 (3), pp. 391-407.

Grafton R. Q., Kompas T., To H. & Ward M/, "Residential Water Consumption: A Cross Country Analysis", Research Report No. 23, Environmental Economics Research Hub Research Reports, The Crawford School of Economics and Government, Australian National University.

Grubb M.D. & Osborns M [2015], "Cellular Service Demand: Biased Beliefs, Learning, and Bill Shock", American Economic Review, 105, pp. 234-71.

Hoover E.M. [1941], "Interstate redistribution of population", 1850-1940, The Journal of Economic History, 1 (2), pp. 199-205.

Howard G & Bartram J [2003], "Domestic Water Quantity, Service Level and Health", World Health Organization, working paper nº 03.02.

Insee Analyses La Réunion [2023], À La Réunion, une même exposition aux risques et nuisances, quel que soit le niveau de vie, n° 84, Octobre 2023

Ito K. [2014], "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing.", American Economic Review, 104 (2), pp. 537-63.

Jessoe K. & Raspon D. [2014], "Knowledge Is (Less) Power: Experimental Evidence from Residential Energy Use", American Economic Review, vol 104, 4, pp. 1417-38.

Komives K., Foster V., Halpern J. & Wodon Q. [2005], Water, Electricity and the Poor - Who benefits from Utility Subsidies?, World Bank, Washington, DC.

Komives K., Halpern J., Foster V., Wodon Q. & Abdullah R. [2007], "Utility subsidies as social transfers: an empirical evaluation of targeting performance", Development Policy Review, 25, pp. 659-697.

Lambert Peter J. [2001], The distribution and redistribution of Income, Manchester University Press, Third Edition.

Lambert P.J. & Aronson J.R. [1993], "Inequality decomposition analysis and the Gini coefficient revisited", Economic Journal, vol 103, pp. 1221-1227.

Lerman R.I. & Yitzhaki S. [1984], "A note on the calculation and interpretation of the Gini index", Economics Letters, vol 15, pp. 363--368.

Liebman J.B. & Zeckhauser R. [2004], "Schmeduling", unpublished manuscript, Harvard University .[75]

[75] www.hks.harvard.edu/jeffreyliebman/schmeduling.pdf

Makki A.A., Stewart R.A., Beal, C.D., Panuwatwanich, K. [2015], "Novel bottom-up urban water demand forecasting model: Revealing the determinants, drivers and predictors of residential indoor end-use consumption", Resources, Conservation & Recycling, 95, pp. 15-37.

March H. C. & Pujol D. S. [2009], "What lies behind domestic water use? A review essay on the drivers of domestic water consumption", Boletín de la Asociación de Geógrafos españoles (A.G.E.), 50.

Martínez-Espiñeira R. & Nauges C. [2004], "Is really all domestic water consumption sensitive to price control? An empirical analysis", Applied Economics, 36 (9), pp. 1697-1703.

Martins R., Cruz L., Barata E. & Quintal C. [2013], "Assessing social concerns in water tariffs", Water Policy 15 pp. 193-211.

Martins R/ & Fortunato A [2007] "Residential water demand under block rates - a Portuguese case", Water Policy 9 (2), pp. 217-230.

Marzano, R., Rouge, C., Garrone, P., Grilli, L., Harou, J. J. & Pulido-Velazquez, M. [2018], "Determinants of the price response to residential water tariffs: Meta-analysis and beyond", Environmental Modelling & Software 101, pp. 236-248.

Mas-Colell A. Whinston M. D. & Green J. R. [1995], Microeconomic Theory, Oxford University Press, New York.

Massey D.S. & Denton N.A. [1988], "The dimensions of residential segregation", Social Forces, 67 (2), pp. 281-315.

Mayol A. [2017], "Social and Nonlinear Tariffs on Drinking Water: cui bono? Empirical Evidence from a Natural Experiment in France", Revue d'économie politique, vol. 127, pp. 1161-1185.

Mayol A. & Porcher S. [2019], "Discriminatory tariffs and drinking water monopolies: an analysis of consumer reaction to price-signal distortions", Revue économique, vol 70, pp. 461-494.

Montserrat T., Bernardo V. & Mirats-Tur J. [2015], Water Demand Models, technical report D4.3, Effinet Project, DOI: 10.13140/RG.2.2.14291.17443.

Morill R.L. [1991], "On the measure of geographic segregation", Geography Research Forum, 11, pp. 25-36.

Mussard S. [2006], "Une nouvelle décomposition de la mesure de Gini en sources de revenu et la décomposition en sous-populations : une réconciliation", Annales d'Économie et de Statistique, 8, pp. 1-25.

Nauges, C. & Thomas, A. [2003], "Long-run study of residential water consumption", Environmental and Resource Economics 26 (1), pp. 25-43.

Nauges C. & Whittington D. [2017], "Evaluating the Performance of Alternative Municipal Water Tariff Designs: Quantifying the Tradeoffs between Equity, Economic Efficiency, and Cost Recovery", World Development, 91, pp. 125-143.

Nordin J. A. [1976], "A proposed modification of Taylor's demand analysis: Comment", The Bell Journal of Economics 7(2), pp. 719-721.

OECD [2010], Pricing water resources and water sanitation services, Technical report, Paris.

Office de l'Eau Réunion [2019], Caractérisation socio-économique des usages de l'eau du Bassin Réunion, Etat des lieux 2019, with Comite de l'Eau et de la Biodiversité - La Réunion & DEAL Réunion,

Paul M. [2023], "Progressive Pricing, Nudges and Economic Efficiency", International Water Week, March 2023, Oviedo.

Pérez Urdiales M., Libra J. M., Machado K., Serebrisky T. & Solís, B. [2022]. "Water Bill Perception in Brazil: Do Households Get It Right?", IDB working paper series, 1336, Inter-American Develoment Bank, .

Pérez-Urdiales M. & García-Valiñas M.A. [2016], "Efficient water-using technologies and habits: A disaggregated analysis in the water sector", Ecological economics, 128, pp. 117-129.

Pérez-Urdiales M., García-Valiñas M.A. & Martínez-Espiñeira R. [2016], "Responses to Changes in Domestic Water Tariff Structures: A Latent Class Analysis on Household-Level Data from Granada, Spain", Environmental and Resource Economics, 63, pp. 167-191.

Porcher S. [2014], "Efficiency and equity in two-part tariffs: the case of residential water rates", Applied Economics, 46, (5), pp. 539-555.

Puri R. & Maas A. [2020], "Evaluating the Sensitivity of Residential Water Demand Estimation to Model Specification and Instrument Choices", Water Resources Research, 56, 1-14.

Rees-Jones A. & Taubinsky D. [2020], "Measuring "Schmeduling"", Review of Economic Studies, 87, pp. 2399-2438.

Reynaud A. [2015], Modelling Household Water Demand in Europe, Insights from a Cross-Country Econometric Analysis of EU-28 countries, JRC Technical Report (27310), European Commission.

Reynaud A. [2008], Social policies and private sector participation in water supply - the case of France, in Prasad Naren (Ed.), Social Policies and Private Sector Participation in Water Supply. Palgrave, Basingstoke & New York, pp. 37-69.

Rinaudo J.D., Neverre N. & Montginoul M. [2012], Simulating the Impact of Pricing Policies on Residential Water Demand: A Southern France Case Study. Water Resources Management, 26, pp. 2057-2068.

Sebri M. [2014], "A meta-analysis of residential water demand studies", Environment, Development and Sustainability 16 (3), pp. 499-520.

Sen A. [1976], "Poverty: an ordinal approach to measurement", Econometrica, vol 44, pp. 219-231.

Schleich J. & Hillenbrand T. [2009], "Determinants of residential water demand in Germany", Ecological Economics, 68 (6), pp. 1756-1769.

Shin J.-S. [1985], "Perception of price when price information is costly: Evidence from residential electricity demand", The Review of Economics and Statistics, 67 (4), pp. 591-598.

Suárez-Varela M., Martínez-Espineira R. & Gonzalez-Gomez F. [2015], "An analysis of the price escalation of non-linear water tariffs for domestic uses in Spain", Utilities Policy, 34, pp. 82-93.

Taylor L. [1975], "The demand for electricity: A survey", The Bell Journal of Economics 6 (1), pp. 74-110.

Theil H. [1972], Statistical Decomposition Analysis, North-Holland, Amsterdam.

Theil H. et Finezza A.J. [1971], "A note on the measurement of racial integration of schools by means of informational concepts", Journal of Mathematical Sociology, 1 (2), pp. 187-194.

Whittington D & Nauges C [2020], "An Assessment of the Widespread Use of Increasing Block Tariffs in the Municipal Water Supply Sector", Oxford Research Encyclopedia of Global Public Health.

Worthington A. C. & Hoffman M. [2008], "An Empirical Survey of Residential Water Demand Modelling", Journal of Economic Surveys (2008) vol. 22, 5, pp. 842-871.

Who - Unicef [2008], Progress on Drinking Water and Sanitation. Special Focus on Sanitation. Unicef / Who, New York and Geneva.

Wilson Robert [1993], Nonlinear Pricing, Oxford University Press, New York

Wichman C.J. [2014], "Perceived price in residential water demand: Evidence from a natural experiment", Journal of Economic Behavior & Organization, 107, pp. 308-323.

Yitzhaki S. [1998], "More than a dozen alternative ways of spelling Gini, Research on Economic Inequality, vol. 8, pp. 13-30.

INN WATER

INN WATER

Promoting social innovation to renew multi-level and cross sector water governance

@InnWater

@ InnWater_EU

https:/www.innwater.eu/

contact.innwater@oieau.fr