IMT School for Advanced Studies Lucca
Lucca, Italy

KU Leuven, Faculty of Economics and Business
Leuven, Belgium

Sub-national Service Provision and Public Spending Analysis

Joint degree in "PhD in Institutions, Markets and Technologies - Curriculum in Economics" and the Doctoral Program in "Business Economics"

XXX Cycle

By

Giovanna D'Inverno

2018

The dissertation of Giovanna D'Inverno is approved.

Program Coordinator:
Prof. Dr. Pietro Pietrini, IMT School for Advanced Studies Lucca

Supervisors:
Dr. Alessandro Belmonte, IMT School for Advanced Studies Lucca
Prof. Dr. Kristof De Witte, KU Leuven, Maastricht University
Prof. Dr. Laura Carosi, University of Pisa

Assessment Committee Faculty Members:
Prof. Dr. Massimo Riccaboni, IMT School for Advanced Studies Lucca
Prof. Dr. Tom Van Puyenbroeck, KU Leuven

The dissertation of Giovanna D'Inverno has been reviewed by:

Prof. Dr. Ana Camanho, Porto University
Prof. Dr. Benny Geys, BI Norwegian Business School

IMT School for Advanced Studies Lucca
KU Leuven, Faculty of Economics and Business

2018

To Filomena, Pasquale and Alessio

Contents

List of Figures **xi**

List of Tables **xiii**

Acknowledgements **xvii**

Vita and Publications **xxi**

Abstract **xxv**

1 Introduction **1**

1.1 Problem statement 3

1.2 Methodological approach 8

1.3 Institutional context 14

1.4 Outline of the dissertation 18

2 Service Level Provision in Municipalities: A Flexible Directional Distance Composite Indicator **25**

2.1 Introduction 25

2.2 Municipal service level 28

2.3 Methodology 34

2.3.1 The BoD model 35

2.3.2 The directional distance BoD model 36

2.3.3 The directional distance BoD model including weight restrictions 38

2.3.4 The robust and conditional directional distance BoD model 40

2.4 Results 43
2.4.1 The municipal service provision composite indicator results 44
2.4.2 Municipal service provision and government size . 49
2.4.3 Robustness checks 51
2.5 Conclusion 51

3 Water pollution in wastewater treatment plants: an efficiency analysis with undesirable output 53
3.1 Introduction 53
3.2 Related literature 56
3.3 Methodology 59
3.3.1 The WWTP production technology 59
3.3.2 The Analytic Hierarchy Process/Non-radial Directional Distance Function approach 61
3.3.3 Identifying WWTP efficiency explanatory variables 64
3.4 Data 65
3.5 The WWTPs performance assessment 67
3.5.1 Model set-up 67
3.5.2 Results 68
3.6 A step further in the WWTPs performance assessment . . 70
3.6.1 External factor choice 70
3.6.2 Results and policy remarks 71
3.7 Conclusion 76

4 Global public spending efficiency in Tuscan municipalities 81
4.1 Introduction 81
4.2 Literature review 84
4.3 Model specification 87
4.3.1 Stage 1: DEA for each individual municipal function 87
4.3.2 Stage 2: Aggregating for overall efficiency analysis 88
4.3.3 Stage 3: Investigating municipal expenditure efficiency explanatory variables 91
4.4 Empirical application 92
4.4.1 Data, inputs and outputs 92

4.4.2 Results 95
4.4.3 A different way of aggregation 99
4.4.4 Efficiency explanatory variables: Tobit regression . 105
4.5 Conclusions 107

5 School Infrastructure Spending and Educational Outcomes in Northern Italy 111
5.1 Introduction 111
5.2 The Italian school system framework 115
5.3 The Northern Italy earthquake overview 118
5.3.1 The May 2012 earthquake 118
5.3.2 The damage evaluation 118
5.3.3 Perception of the risk, reconstruction and funding . 119
5.3.4 Government actions after the earthquake 121
5.4 Data sources and summary statistics 122
5.4.1 Educational achievement and school data 122
5.4.2 Earthquake damage assessment and seismic hazard 123
5.4.3 School funding data 124
5.5 Empirical strategy 126
5.5.1 Difference-in-differences estimation strategy 130
5.5.2 Instrumental variable strategy 132
5.6 Results 132
5.7 Conclusion 138

6 Conclusions and Discussion 141
6.1 Evidence-based policy implications 145
6.2 Study limitations and future lines of research 149

A Supplementary material for Chapter 2 153
A.1 Data sources and description 153
A.2 Municipal-specific weights 155
A.3 Spearman correlation tables 167
A.4 Shadow prices results 169
A.5 Statistical inference results 172
A.6 Government size results 174

A.7 Robustness check results 175

B Supplementary material for Chapter 3 **179**
B.1 Some essential notions of wastewater treatment 179
B.2 Water Performance Indexes results 182

C Supplementary material for Chapter 4 **185**
C.1 Municipal efficiency scores and composite indicators at local and global level 185

References **191**

List of Figures

1.1 Full list of the 17 proposed Sustainable Development Goals 3

2.1 Example of services for each municipal function 29
2.2 Intertemporal variation of the service provision when considering economic variables 47
2.3 "Inverted U-relationship" between municipal tax burden and municipal service provision level 50

4.1 Geographical distribution comparison. 97
4.2 Comparison between composite indicators scores by dimensional classes 100
4.3 Comparison between composite indicators scores by mountain classes 100
4.4 Comparison between composite indicators scores by tourism classes 100
4.5 Comparison between composite indicators scores by local labour system classes 101
4.6 Municipalities by relative efficiency and expenditure composition. 103

5.1 Financial flow diagram for the general upper secondary schools 117
5.2 Epicentral area affected by the seismic events 119
5.3 Buildings vulnerability and damage level. 120

5.4 Average number of enrolled students (treated and control group) 128
5.5 Treated and control areas 129
5.6 Estimated impact of school funding on test scores by quantiles of the distribution of test scores 137
5.7 Estimated impact of receiving extra-funding on test scores by quantiles of the distribution of test scores 138

A.1 "Inverted U-relationship" between municipal tax burden and municipal service provision level 174
A.2 "Inverted U-relationship" between municipal tax burden and municipal service provision level (Analysis without "Centrumsteden") 177

List of Tables

1.1 Dissertation outline 22

2.1 Descriptive statistics for the municipal service level 31
2.2 Descriptive statistics for the municipal background conditions 34
2.3 Summary of the weights obtained from the municipal expenditure composition 40
2.4 Descriptive statistics of the service provision composite indicator scores estimated for 307 municipalities over 2006–2011 45
2.5 Influence of background conditions on municipal service composite indicator 48
2.6 Optimal tax rates across different model specifications . . . 50

3.1 Indexes 64
3.2 Descriptive statistics 67
3.3 AHP-non-radial set of weights 68
3.4 Efficiency score for each Wastewater Performance Index . 69
3.5 WPIs by explanatory factors and Mann-Whitney U/Kruskal-Wallis test results 74
3.6 WPIs by explanatory factors and Mann-Whitney U/Kruskal-Wallis test results 75

4.1 Descriptive statistics for DEA dataset – Input. 95
4.2 Descriptive statistics for DEA dataset – Output. 95

4.3 Descriptive statistics of the efficiency scores at local and global level. 96
4.4 Descriptive statistics of global efficiency scores. 98
4.5 Descriptive statistics of global efficiency scores (cont'd). . . 99
4.6 Descriptive statistics of the scores for municipal and Tuscan average weights. 101
4.7 Descriptive statistics of each quadrant. 104
4.8 Tobit results. 107

5.1 Summary Statistics 125
5.2 Pre-treatment Test Scores 130
5.3 Pre-treatment Covariates 130
5.4 Secondary School, Mathematics: funding and students scores 135
5.5 Secondary School, Italian: funding and students scores . . 136

A.1 Downloaded variables used for the empirical application. . 154
A.2 Lower bound municipal-specific weights 155
A.3 Upper bound municipal-specific weights 161
A.4 Spearman correlation among weight restriction models . . 167
A.5 Spearman correlation among model specifications 168
A.6 Descriptive statistics of the shadow prices for each municipal function (Unrestricted unconditional and Robust conditional estimates) 170
A.7 Descriptive statistics of the budget shares across the municipal functions (Unrestricted unconditional and Robust conditional estimates) 171
A.8 Influence of background conditions on municipal service composite indicator 173
A.9 Optimal tax rates across different model specifications . . . 174
A.10 Descriptive statistics of the service provision composite indicator scores estimated for 294 municipalities over 2006–2011 (Analysis without "Centrumsteden") 175
A.11 Influence of background conditions on municipal service composite indicator (Analysis without "Centrumsteden") . 176

B.1 Indexes 182
B.2 Water efficiency Performance Indexes results 183

C.1 Municipal efficiency scores and composite indicators at local and global level 185

Acknowledgements

Chapter 2 is the reproduction of the paper "Service Level Provision in Municipalities: A Flexible Directional Distance Composite Indicator", co-authored with Kristof De Witte.

Chapter 3 is the reproduction of the paper "Water pollution in wastewater treatment plants: an efficiency analysis with undesirable output", "*European Journal of Operational Research*" (2017). The paper is co-authored with Laura Carosi, Giulia Romano and Andrea Guerrini.

Chapter 4 is the reproduction of the paper "Global public spending efficiency in Tuscan municipalities", "*Socio-Economic Planning Sciences*" (2017). The paper is co-authored with Laura Carosi and Letizia Ravagli.

Chapter 5 is based on the paper "School Infrastructure Spending and Educational Outcomes in Northern Italy", *SPRU Working Paper Series (SWPS)* (2017). The paper is co-authored with Alessandro Belmonte, Vincenzo Bove and Marco Modica.

This doctoral dissertation would not have been possible without the advice and support of many people. First and foremost, I would like to express my deep gratitude to my supervisors, *Dr. Alessandro Belmonte*, *Prof. Dr. Kristof De Witte* and *Prof. Dr. Laura Carosi*. Thank you for sharing your knowledge with me, for your valuable advices and your continuous guidance throughout this journey, for providing me all I needed to successfully complete my PhD, for always believing in me and in my abilities. *Alessandro*, thank you for giving me the opportunity to work on a very interesting project, for our fruitful discussions and for your help. *Kristof*, thank you for accepting me as visiting student and for offering me so many opportunities since I joined your research group. *Laura*, thank you for your assistance since my very first year of bachelor, almost ten years ago, for your understanding and ceaseless encouragement.
I would also like to thank the external referees *Prof. Dr. Ana Camanho* and *Prof. Dr. Benny Geys*, together with the other committee members,

Prof. Dr. Tom Van Puyenbroeck and *Prof. Dr. Massimo Riccaboni*, for the invested time and effort in reading my dissertation. I am very grateful for your insightful comments and constructive suggestions that improved the overall quality of my dissertation. I also appreciate very much your willingness to travel and to attend my close and public PhD defence.
My sincere gratitude goes to my co-authors, *Dr. Giulia Romano*, *Prof. Dr. Andrea Guerrini*, *Letizia Ravagli*, *Prof. Dr. Vincenzo Bove* and *Dr. Marco Modica*. Thank you for your contribution and collaboration, for your kind support and motivation during these years working together. I would also like to thank every researcher and professor who have contributed to my work and my personal growth as an academic. In particular, I would like to thank *Prof. Dr. Wim Moesen*, *Prof. Dr. Riccardo Cambini* and *Prof. Dr. Mauro Sodini* for their enthusiasm and passion for research and for teaching, for their constant positive and motivating attitude, for their valuable advices and encouragements. I am deeply grateful to *Serena Argentieri*, *Daniela Giorgetti* and *Sara Olson* in Lucca and *Ann Van Espen* in Leuven for their excellent administrative support during my PhD. Special thanks to all those who made possible the joint PhD at IMT and at KU Leuven.

These last years have been fundamental in my life and I am glad so many people have been part of it. In particular, I would like to thank my colleagues in Lucca, *Luca*, *Armando* and *Loredana*: thanks for all the challenges we faced together and for the mutual support through all the deadlines and the assignments. Thanks to my present and past colleagues in Leuven, *Fritz*, *Deni*, *Vita*, *Kenneth*, *Ruben*, *Joana*, *Kaat* and *Boukje*: thanks for welcoming me in an extremely nice way, for our mutual esteem, for your enthusiasm for my research since the very first time we met each other, for all the tips you shared with me.
I would like to thank IMT for giving me the chance to meet many students from all over the world. You all, guys, gave me the opportunity to get a glimpse of the different cultures spread around the world and to travel through your stories, traditions and life experiences, in China, in India, in Iran, in Turkey, in Ukraine, in Russia, in Germany, in UK and in Argentina. In addition to this, many thanks to the IMT Lucca commu-

nity for the crazy evenings in the living room, the late nights studying and eating popcorn and more than I could write in few lines. Thank you all for all the memorable moments we shared together. Thank you *Ele* for our roomie late night chats, *Pakhee* and *Lucia* for your spell check support and our days in the study room. Special thanks to *Valentina*, for patiently listening to my mock presentations before almost every conference I attended, for giving me straight suggestions without sugaring the pill, for making many situations seem less daunting. Thank you *mere pyare doston Vihang*, *Manas* and *Soumali*, for your extreme friendliness, many masala chai and Indian dinners. Thank you *Viggy* for motivating me all the time and making sure I could overcome any situation. Thank you *Ilkay*, *Unal*, *Anita*, *Vincenzo*, *Daria*, *Mika*, *Vitaly*, *Paolo*, *Alessandro*, *Valerio* Palermo and *Valerio* Catania for the wonderful time. Many thanks also to the KU Leuven community, for your warm welcome and letting me feel at home in Belgium. A special thank to *Puya* and *Andreas*, my guardian angels since the very first day I arrived in Leuven, *Prosci* and *Botti*, *Presijdent* and *his crew*, "*piccoli*" and "*piccole*".
Furthermore, I would like to thank all my family, grandparents, uncles and cousins, and all my old and new friends for their support and affection throughout my study. In particular, thanks to *Sara*, *Gloria*, *Maria*, *Lorenzo*, the *original "cocche" Camilla* and *Giulia*, *Lucia*, *Naomi* and *Valeria*, *Rita* and *Erika*, *Sara* and *Silvia*, all the "*cocchi e non*" around Viareggio, Pisa and Sarzana, my dearest friends that have always been beside me in the difficult moments urging me not to be discouraged and also shared many milestones during these years. My deep gratitude goes to "*Corradis B&B*": I cannot imagine how it could have been possible to travel from Italy to Belgium and back in such a comfortable way without your priceless hospitality. I would also like to thank my flatmates for their kindness, especially when deadlines were approaching: *Liesbeth*, *Lina* and *Arianna*, *Agnese*. Thanks to *Gabriella*, for her continuous appreciation and thoughtful company in Leuven. Thanks to the support of all the guys from the church of Grosseto, Pisa and beyond, till Bruxelles: thank you *Cettina* and *Biagio* for giving me the warmth of a family despite being far from home.

I would like to express my deepest gratitude to my mum *Filomena*, my dad *Pasquale* and my brother *Alessio*, who have never let me feel alone in my decisions and granting me their unconditional love all the time. I know that no matter what you will be always there for me.
Above all, I am grateful to the *Lord*, who made all this possible, in a way that I could have never imagined, who keeps my life in His safe hands and led me up to this point, because

> "*He giveth power to the faint; and to them that have no might he increaseth strength. Even the youths shall faint and be weary, and the young men shall utterly fall: But they that wait upon the Lord shall renew their strength; they shall mount up with wings as eagles; they shall run, and not be weary; and they shall walk, and not faint.*"
>
> (Isaiah 40:29–31)

Vita

December 31, 1989	Born, Grosseto, Italy

Education

2017–	Ph.D. candidate jointly in Business Economics Faculty of Economics and Business, KU Leuven, Belgium
2014–	Ph.D. candidate in Economics, Management and Data Science (XXX cycle) IMT School for Advanced Studies Lucca, Italy
2013	M.Sc. in “Economics” Final mark: 110/110 cum laude, Excellence Path University of Pisa and Scuola Superiore Sant’Anna, Italy
2011	B.Sc. in “Economia e Commercio” Final mark: 110/110 cum laude, Excellence Path University of Pisa, Pisa, Italy
2009	Piano Diploma Final mark: 10/10 cum laude Conservatorio “L. Perosi”, Campobasso, Italy

Research Activity

10/2017–	“Steunpunt voor Onderwijsonderzoek SONO” project Leuven Economics of Education Research Faculty of Economics and Business, KU Leuven, Belgium
01/2017–09/2017	Visiting at Leuven Economics of Education Research Faculty of Economics and Business, KU Leuven, Belgium

02/2014–09/2014	"Spending Review on the Healthcare Personnel Count" Istituto Regionale Programmazione Economica della Toscana (IRPET), Firenze (FI), Italy
07/2013–09/2013	Internship Istituto Regionale Programmazione Economica della Toscana (IRPET), Firenze (FI), Italy

Working Experience

09/2017–	Contract professor for the Mathematics undergraduate course (Fall 2017) University of Pisa, Italy
10/2014–09/2017	Teaching Assistant for the Mathematics undergraduate course (Fall 2014–2016) University of Pisa, Italy

Educational Activities

2017	EdEN Scientific Workshop Leuven (Belgium), March 29, 2017
2016	PhD Symposium on "Efficiency in Education-concepts, techniques and results" Milan (Italy), October 17-19, 2016
2016	Barcelona GSE Summer School 2016, "Quantitative Methods for Public Policy Evaluation" Barcelona (Spain), July 4-8, 2016
2016	EdEN Workshop, "Structural Models for Policy Evaluation and Discrete Choice Approaches" Maastricht University & TIER, Maastricht (The Netherlands), June 6-8, 2016

Honors, Grants and Awards

1. AMASES Award for the best paper presented by a young researcher, at XLI National AMASES Conference, Cagliari (Italy), September, 2017
2. ERASMUS+ MOBILITY CONSORTIUM Grant (Academic Year 2016/2017)
3. Ph.D. Full Scholarship in Economics (XXX cycle), IMT - School for Advanced Studies, Lucca, Italy
4. Ph.D. Full Scholarship in Analitics for Economics and Business (XXX cycle), Universit degli studi di Bergamo, Italy
5. M.Sc degree with Honors
6. B.Sc degree with Honors

Publications

1. Giovanna D'Inverno, Laura Carosi, Letizia Ravagli, 2017. "Global public spending efficiency in Tuscan municipalities", in *Socio-Economic Planning Sciences*
   `http://dx.doi.org/10.1016/j.seps.2017.01.006`
2. Giovanna D'Inverno, Laura Carosi, Giulia Romano, Andrea Guerrini, 2017. "Water pollution in wastewater treatment plants: An efficiency analysis with undesirable output", in *European Journal of Operational Research*
   `https://doi.org/10.1016/j.ejor.2017.08.028`
3. Alessandro Belmonte, Vincenzo Bove, Giovanna D'Inverno, Marco Modica, 2017. "School Infrastructure Spending and Educational Outcomes in Northern Italy" in *SPRU Working Paper Series (SWPS)*, 2017-20: 1-23. ISSN 2057-6668

Presentations

1. G. D'Inverno, K. De Witte, "Municipal Performance and Government Size: Evidence from Flanders", at *5th Workshop on Efficiency in Education–"Efficiency in education and performance measurement in public sector"*, Budapest (Hungary), October 19–20, 2017
2. G. D'Inverno, L. Carosi, L. Ravagli, "Global public spending efficiency in Tuscan municipalities", at *XLI National AMASES Conference*, Cagliari (Italy), September 14–16, 2017
3. A. Belmonte, V. Bove, G. D'Inverno, M. Modica, "School Infrastructure Spending and Educational Outcomes in Northern Italy", Invited seminar at *Faculty of Economics and Business*, KU Leuven, Leuven (Belgium), June 22, 2017
4. A. Belmonte, V. Bove, G. D'Inverno, M. Modica, "School Infrastructure Spending and Educational Outcomes in Northern Italy", Invited seminar at *Top Institute for Evidence-Based Education Research (TIER)*, Maastricht (The Netherlands), June 20, 2017
5. L. Carosi, G. D'Inverno, G. Romano, A. Guerrini, "Estimating efficiency of Italian water utilities by accounting for quality issues", at *15th European Workshop on Efficiency and Productivity (EWEPA) 2017*, London (UK), June 12–15, 2017
6. G. D'Inverno, K. De Witte, "Municipal Performance and Government Size: Evidence from Flanders", at *EdEN Summer School*, Leuven (Belgium), May 29–31, 2017
7. A. Belmonte, V. Bove, G. D'Inverno, M. Modica, "School Spending and Students' Achievement: Evidence from the Northern Italy Earthquake", at *3rd Workshop on Educations Economics*, Leuven (Belgium), March 30–31, 2017
8. A. Belmonte, V. Bove, G. D'Inverno, M. Modica, "School Spending and Students' Achievement: Evidence from the Northern Italy Earthquake", at *4th Workshop on the Efficiency in Education*, Milan (Italy), October 18–19, 2016
9. G. D'Inverno, L. Carosi, L. Ravagli, "Global public spending efficiency in Tuscan municipalities", at *AIRO 2015 CONFERENCE*, Pisa (Italy), September 7–10, 2015
10. G. D'Inverno, L. Carosi, L. Ravagli, "Global public spending efficiency in Tuscan municipalities", at *DEA2015*, Braunschweig (Germany), August 24–27, 2015

Abstract

Local and regional governments represent the nearest form of government to the people and their fundamental role in addressing citizens' needs is acknowledged worldwide. In line with the subsidiarity principle, the responsibility of the public good and service delivery is primarily upon the territorial administrations closer to the citizens, with the main advantage of offering more suitable and better tailored solutions at local level. In a context of scarce resources and tight budget constraints exacerbated by the ongoing economic crisis, the achievement of these goals is limited and needs to be pursued in an efficient and effective way. Accordingly, the development of tools to evaluate the performance of local and regional government is required, as well as measures to monitor the progress of the task achievement and instruments to support over time the decisional process, in the interest of all the involved shareholders, specially policy makers and citizens, international and civil society organizations.

This dissertation contributes to the knowledge on basic service delivery and public expenditure analysis at sub-national level. Particularly, it deals with the provision of both general and specific services, namely the education and water sector ones. From a methodological point of view, innovative methods are proposed to evaluate the service supply and public spending in terms of efficiency and effectiveness. To show the potentiality of the suggested tools, empirical applications are proposed covering two EU countries, Belgium and Italy, which are interesting study cases for their common and peculiar features and provide complementary insights.

In Chapter 2, we propose the innovative use of a composite indicator to measure the multidimensional aspects of the local public provision, encompassing several commonly acknowledged municipal tasks, and to investigate the relationship with the local government size, as the decentralization of public activities to the municipalities calls for a more enhanced service provision analysis at the local level. We suggest a robust conditional version of a directional distance Benefit-of-the-Doubt approach with weight restrictions based on the municipal expenditure composition. Specifically, we deal with the presence of undesirable municipal service indicators and with the heterogeneity among the municipalities in their political preferences, priority public activities and operating environment characteristics. To illustrate the applicability of the suggested method, we show the construction of the municipal service provision composite indicator for 307 Flemish municipalities over the year 2006-2011.

As a focus on a particular service, in Chapter 3 the environmental efficiency of 96 Tuscan (Italian) wastewater treatment plants (WWTPs) is investigated taking into account the quality of the outgoing water in terms of pollutant. In this regard, the presence of the residual nitrogen in the outgoing treated water is considered as undesirable output. The efficiency analysis is performed by applying a novel integrated AHP/non-radial directional distance function approach. The obtained results are then used to identify the efficiency explanatory variables: among them, the facilities' capacity, the percentage of wastewater discharged by the industrial and agricultural activities and the level of compliance with the pollutant concentration threshold set by the legislator have a significant impact on the WWTP performance.
In Chapter 4, a Data Envelopment Analysis model is used to study the efficiency of Tuscan municipalities' public expenditure. Five strategic functions of Tuscan municipalities

are first considered carrying out a non-aggregate analysis; then the overall expenditure composition of each municipality and the global spending efficiency are analysed by a proposed composite indicator. The main determinants affecting the municipalities' efficiency are further investigated. In particular, the obtained results may be consistently included in the long-standing debate on the municipal size, proving that the bigger the municipality, the greater its level of public expenditure efficiency.

In Chapter 5, we explore whether investment in public school infrastructure affects students' achievement. We use data on extra funding to public high schools after the 2012 Northern Italy earthquake and apply a quasi-experimental design and an instrumental variable strategy. We find that spending on school infrastructure increases standardized test scores in mathematics and Italian language, and the effect is stronger for lower-achieving students and in mathematics. These results provide evidence in favour of a positive impact of capital spending in improving the learning environment and performances of high school students.

Chapter 1

Introduction

Since 2015, countries from all over the world have adopted the '2030 Agenda for Sustainable Development' to ensure prosperity, social inclusion and environmental sustainability for the benefit of people, planet, prosperity, peace and partnership (UN General Assembly, 2015). The Agenda consists of 17 Sustainable Development Goals (SDGs), 169 targets and 240 indicators to be achieved by 2030. "The SDGs cover issues that are directly relevant to citizens' daily lives, including vital challenges such as poverty, gender inequality, climate change and insecurity, as well as public goods like education, health, water, energy, air quality, housing and the conservation of natural resources" (Global Taskforce of Local and Regional Governments, 2016). In compliance with the subsidiarity principle, these activities should be provided by the territorial administration closer to the citizens. The rationale behind this principle lies in the fact that the forms of government more proximate to the people have the chance to better identify and fulfil the citizens' need (Slack and Bird, 2013). They can offer more tailored solutions and enable citizens' preferences to be taken into account more accurately, pursuing a higher level of efficiency in the expenditure management (Da Cruz and Marques, 2014). Furthermore, a closer interplay between local authorities and locals aims at promoting higher responsibility and participation at local level, increasing the level of accountability and stimulating forms of local competition (Asatryan and De Witte, 2015). The local and regional governments represent the form of government closest to the people and their fundamental role in addressing citizens' needs is recognized worldwide (OECD, 2016a). Accordingly, sub-national governments play a crucial coordination part while fostering citizen participation and bringing higher forms of government closer to the people for the SDGs attainment.

While playing their essential role in delivering basic services, local and regional governments have to operate efficiently and effectively in a context of scarce resources and tight budget constraints, which became even more stringent after the global economic crisis started in 2008. As a consequence, monitoring the basic service provision and, more broadly speaking, the SDGs implementation becomes essential: "public sector organisations, departments and agencies regularly monitor user and citizen satisfaction with public services to evaluate the impact of reforms and identify areas calling for further actions" (OECD, 2017c). This helps in supporting and improving over time the performance of local and regional governments, which should collaborate to ensure a more efficient and integrated approach in the territorial development cooperating in public good and infrastructure delivery, sharing skills and resources (Global Taskforce of Local and Regional Governments, 2016). Tools apt to evaluate the service provision and to control expenditure need to be developed in the interest of all the involved shareholders, among all policy makers and citizens, international and civil society organizations. The achievement of these objectives represents a necessary step towards an effective, innovative and sustainable financing system (UN General Assembly, 2017).

This dissertation contributes to the knowledge on basic service delivery and public expenditure analysis at sub-national level. Specifically, it proposes innovative methodologies to address three SDGs-related issues: education facility upgrade for effective learning environment as part of Goal 4, water sanitation as part of Goal 6, making cities inclusive and sustainable as part of Goal 11 (for the full list of SDGs we refer to Figure 1.1). To show the potentiality of the tools and the techniques suggested in this dissertation, empirical case studies are proposed covering

Figure 1.1: Full list of the 17 proposed Sustainable Development Goals
Source: https://unwomen.org.au/

two EU countries, Belgium and Italy.

In the following, we first introduce more in depth the problem statements addressed in this dissertation. Then, we give a sketch of the theoretical framework on which we rely to give answers to the raised research questions: we show how we make a step forward compared to the existing literature by proposing the innovative use of enhanced techniques. Furthermore, we explore the institutional background on which we frame our analysis, giving emphasis to the distinctive features of the contexts under analysis. Finally, the outline of the dissertation is presented.

1.1 Problem statement

In this dissertation, we contribute to the knowledge on basic service provision and public spending measurement at sub-national level, considering methodological and empirical issues. These two topics are complementary and sides of the same coin. From the one hand, the key for a sustainable development originates at local level providing services closer and more tailored to citizens' need and promoting a bottom up process. On the other hand, tight and scarce resources limit public goods provision, forcing the local budget to be efficiently and effectively aligned with the priorities defined by the political preferences and the specific features of the local context. Accordingly, the development of tools to evaluate the performance of local and regional government is required, as well as measures to monitor the progress of the task achievement and instruments to support over time the decisional process.

Peer-learning and benchmarking are unanimously recognized as effective ways to improve government performance (Da Cruz and Marques, 2014). Local and regional tasks cover a broad variety of intervention areas, such as educational and social care services, water sanitation and waste disposal management, local security and housing services, among others. The construction of a composite indicator turns out to be a useful tool to encompass all these aspects in a single index (Karagiannis, 2017; Nardo et al., 2005; OECD, 2008), enabling then the comparison among the observed units. However, as concerns the estimation of a synthetic score for service delivery measurement, a number of issues has to be taken into account and multifaceted aspects have to be considered. In fact, local government competencies reflect a multidimensional framework and a significant degree of heterogeneity pervades different spheres of action. For example, municipalities vary in their political preferences and in their priority activities. Yet, sub-indicators measuring municipal undesirable features, such as the local crime level the local police has to deal with, deserve to be modelled carefully and at the same time also the operating context needs to be handled in a proper way. Additionally, from a policy perspective the interest in measuring the overall level of delivered services is relevant not only to detect the best practices, but also to get insights relating the taxation imposed on citizenship, as citizens pay taxes for the services they receive. The presence of a trade-off between the amount of provided public goods and the resources necessary to produce them has been long questioned in the economic literature. Specifically, there is a significant literature investigating the existence of an optimal government size across countries (see for example De Witte and Moesen, 2010), but not across municipalities so far: this kind of analysis would investigate whether the level of tax burden is fair given the overall level of services or whether there is room to change the local current government size, looking at the other realities. Chapter 2 adds to the existing literature providing an innovative tool apt to include all the raised concerns and suitable to further explore the issue of citizens' tax burden.

The overall municipal service provision analysis helps to get a broad overview of the level of public goods delivered at local level. However, each local action deserves to be addressed also more in detail, so to go more in depth and to add complementary aspects for a more insightful picture of the public sector management. In this regard, water management attracts a great deal of attention nowadays. Water is a fundamental component of human life, but it is not accessible to everyone and environmentally sustainable everywhere yet. Moreover, given the huge infrastructural costs related to this industry, an increasing collaboration building on public-private partnerships is taking off. As a possible consequence, the service might be run in compliance with the private sector management criteria, that is mainly in terms of economic and financial profitability. However, when assessing the way water services are provided, their performance efficiency should be evaluated not only in terms of economic profitability, but also in line with the environmental sustainability aspects for the sake of the people and the planet. In this regard, the presence of residual pollutants in the treated outgoing water cannot be neglected and suitable tools should be developed accordingly. The wastewater treatment plants that are more "environmentally" focused should be valued more for keeping their water quality commitment rather than penalized, as they have on the one hand higher costs of production, but on the other hand higher quality outcome. Accordingly, Chapter 3 deals with this topic and provides a novel integrated approach to measure the wastewater treatment plants, encompassing pollutants' issues and including environmental sustainability in the water service provision evaluation.

However, in the public sector management the focus is not only on the service provision side, but also on the way resources are spent. The necessity of combining public service delivery with the containment of public spending has been considered a keyword of the 'New Public Management' paradigm (Hood, 1991). Despite several criticisms to this approach in the last couple of decades (for a detailed discussion, see Hyndman and Lapsley, 2016, and the references therein), the call for efficiency and on effective way of delivering public services still applies. The use of performance evaluation tools and the implementation of an effective system of incentives are in the agenda of both politicians and academics. Therefore, the question that needs to be addressed is how to evaluate the local government service provision encompassing several municipal tasks, while including at the same time the spending efficiency analysis aspects. In line with the subsidiarity principle already mentioned above, local governments are the most involved organizations in this evaluation process, given their increasingly important role in delivering basic services and addressing citizens' needs. One of the most debated sources of spending mismanagement is related to municipal population size. As expressed in a report by the Council of European Municipalities and Regions, "despite the diversity of the municipal level, some general trends are visible at European level as municipalities share common preoccupations. One of their main preoccupations is the quest for the perfect size which would ensure both local democracy and economic efficiency in the delivery of local public services. Different solutions are put in place in order to reach this goal" (Hermenier, 2009). Across Europe, the proposed solutions to attain the perfect size have been mainly either voluntary or mandatory merger policies or inter-municipal cooperation forms. The main difficulty associated to small municipalities is the inability to exploit scale and scope economies. However, the debate is still ongoing and no optimal solution has been found yet. On the one hand, the need to provide more tailored and varied services calls for a joint production. On the other hand, the local identity deserves to be preserved and valued. For these reasons, the development of tools suitable to measure local expenditure efficiency in service provision is fundamental to provide policy makers objective evidence-based recommendations. Chapter 4 contributes to the literature providing an integrated approach in three stages to evaluate both separately and together the main municipal functions' spending efficiency and addressing specifically the municipal size issue.

As outlined above, public expenditure has to be managed in compliance with the principles of both efficiency and effectiveness. Using the words of Peter Drucker, "efficiency is doing things right; effectiveness is doing the right things". Despite the fact that isolating the two concepts might be tricky, we can say that effectiveness analysis aims at showing how successfully resources have been used in reaching the set goals. The funding in the education sector represents a significant expenditure item on the overall level of public spending, as education is considered one of the most important drivers for the long run growth of a country (European Commission, 2014). In particular, school capital funding aims both at constructing new school places and at improving the condition of existing school buildings, by keeping them safe and adequate, reinforcing anti-seismic measures and ameliorating the building system. Nowadays, the lack of investment in school infrastructure is an acknowledged issue worldwide [1]. However, on the one hand the effect of resources and, more broadly speaking, of school spending on students' achievement is still debated. Since the "Coleman Report" (Coleman, 1966), empirical studies provided mixed evidence, showing a lack of agreement. On the other hand, the role of the schools' physical environment and their facilities condition in explaining variation in student learning across schools has been emphasized by educational researchers, social psychologists and sociologists (Bakó-Biró et al., 2012; Earthman, 2002; Haverinen-Shaughnessy et al., 2015; Mendell and Heath, 2004). The main idea is that a school environment offers the students a more comfortable place where to study when it is better-maintained, making the

[1] In the coming years, huge investments in school infrastructure are planned in Australia (https://www.nsw.gov.au), in Italy (http://www.istruzione.it), in India (http://www.cerestraedufund.com), in UK (https://www.nao.org.uk), in US (https://www.washingtonpost.com), in Germany (https://www.ft.com), to name few of them.

learning process more efficient. Chapter 5 investigates this subject and attempts to give an answer to the debated theme, using two intertwined policy evaluation techniques and making use of information obtained after a series of seismic events that occurred in 2012 in Italy.

1.2 Methodological approach

This work aims at proposing innovative tools and policy evaluation techniques that can support all the stakeholders involved in the decision-making process, so to monitor the basic service delivery and to evaluate the resource management, assessing and promoting local action towards the achievement of the set goals.

The efficiency and effectiveness analysis is an important theme in the public sector performance literature. As concerns the efficiency analysis, there are mainly two approaches, parametric and the non-parametric one, that differ in the way the unknown and unobservable "efficiency frontier" is inferred from the data. The first one imposes a specific production function and estimates the error term as the deviation from the given technology consisting in both statistical noise and detected level of inefficiency. The econometric "Stochastic Frontier Analysis" (SFA) introduced by Aigner et al. (1977) represents the leading approach among the parametric techniques. Conversely, the non-parametric approach relies on mathematical programming techniques to evaluate the relative efficiency of one unit compared to the others. The most commonly employed optimization tool is the linear programming model referred to as "Data Envelopment Analysis" (DEA), introduced by Charnes et al. (1978) and based on the concept of efficiency proposed by Farrell (1957). The non-parametric approach is particularly suitable for public sector efficiency analysis: it avoids assuming any specific functional form of the production frontier, as we do not have a priori knowledge of the public sector performance functional form nor understanding of the importance of the different intervention areas, and it gives useful insights to correct the detected level of inefficiency. With a similar mathematical formulation and inspired by the DEA methodology, the Benefit-of-the-Doubt

(BoD) weighting technique has been proposed in the literature as an aggregating method to group several dimensions into one single composite indicator (CI) and labelled as such after Melyn and Moesen (1991). The composite indicator is a tool commonly acknowledged by academics and research to group in a synthetic index a multidimensional phenomenon, not only for benchmarking and performance comparison purposes, but also for policy analysis and public communication (Nardo et al., 2005; OECD, 2008; Zhou et al., 2010). In particular, the peculiarity of the BoD model is that it assigns endogenously the weights for each sub-indicator: the weights are optimally chosen so to give more importance to what they can do the best and low importance to what they can do the worst.

As for the methodological contribution to the literature provided by this work referring to the composite indicator and the efficiency analysis tools, Chapter 2 introduces the development of a new flexible directional distance function composite indicator to measure the municipal service provision; Chapter 3 proposes a novel integrated Analytic Hierarchy Process/Non-radial Directional Distance Function (AHP/NDDF) approach to evaluate the environmental efficiency of wastewater treatment plants service supply, with a particular focus on the undesirable output inclusion in the model specification; Chapter 4 provides a novel 3-stage DEA based approach to get an overall picture of the municipal spending efficiency evaluation.
More specifically, in Chapter 2 we propose a 'dynamic robust conditional directional distance Benefit-of-the-Doubt model with ARI restrictions', encompassing several municipal intervention areas and integrating the main advantages of existing model specifications to address a number of issues that this kind of analysis can bring out. First of all, we deal with heterogeneity among the municipalities in their political preferences and in their priority public activities: we combine the insights of fully non-parametric techniques to grant each municipality the benefit of the doubt (Cherchye et al., 2007; Melyn and Moesen, 1991) with the weight restriction specification to include municipal budget allocation information. We also handle the presence of "undesirable" municipal service indicators by using a directional distance function as proposed by Zanella et al.

(2015), as "more" is not always "better" along the components we evaluate the municipal services. Moreover, we perform the analysis within a robust and conditional approach to include the local government operating environment characteristics (e.g. the income level, the citizens' structure, the political orientation), the time dimension and to correct for outlying observations (Bădin et al., 2012; Daraio and Simar, 2007; Mastromarco and Simar, 2015). The proposed composite indicator is a step forward in the Operational Research literature in several directions. First, there are no studies nor composite indicators encompassing all the listed issues and/or measuring in a comprehensive way the overall municipal service provision (for a review, see Fusco et al., 2017; Karagiannis, 2017); alternatively there are studies focusing only on single services (Rogge et al., 2017) or not including undesirable features (Afonso and Fernandes, 2008; Yusfany, 2015). Second, so far the expenditure composition has not been included directly in the model in the form of weight restriction specification, but only in the municipal task aggregating scheme (Bosch et al., 2012; Helland and Sørensen, 2015). Finally, only few papers deal with the robust and conditional version of the municipal performance evaluation (Asatryan and De Witte, 2015; Cordero et al., 2016).
Moving from an analysis encompassing all the main local government competencies to a specific service sector, a novel integrated Analytic Hierarchy Process/Non-radial Directional Distance Function (AHP/NDDF) approach is presented in Chapter 3, aiming at assessing the wastewater treatment plants in terms of environmental efficiency, including for the first time the presence of residual nitrogen as undesirable output. The benefits of the two techniques are combined and explained as follows. Similarly to the standard NDDF approach (Adler and Volta, 2016; Zhou et al., 2012), the suggested model allows to include simultaneously inputs, desirable and undesirable outputs and not to overestimate the efficiency scores. At the same time, the AHP inclusion gives the possibility to directly take into account the decision maker preferences in the weighting system and to encompass some existing directional distance function models as special cases. Although several integrated Data Envelopment Analysis/Analytic Hierarchy Process models have been pro-

posed in the literature (see for example Pakkar, 2015), the present one addresses different problems and purposes. Specifically, this model considers the sustainability aspect in the wastewater treatment process assessment, to put emphasis on its environmental impact in terms of pollutants left in the outgoing water. The efficiency analysis literature uses the notion of undesirable output, when referring to outputs whose increase might not be desirable. In the water sector performance assessment, only few DEA papers deal with the undesirable output. The notion these studies use has a broader meaning with respect to other fields of application, such as energy and cement sector: in the water service sector literature, the undesirable outputs encompass unintended bad consequences. There are several ways to model undesirable outputs, depending on how the production technology process has been formalized. In compliance with the wastewater treatment plants' activity, we model the undesirable output according to the null-jointness and weak disposability assumptions.

Finally, Chapter 4 combines the insights of the efficiency model and the use of composite indicators, by proposing a novel 3-stage DEA based approach to get an overall picture of local government spending efficiency, integrating the evidences stemming both from a separate and from a global analysis. In the first stage, we run a DEA model for each municipal function under analysis, to explore the expenditure efficiency at non-aggregate level in a fully flexible way. The second stage concerns the aggregate analysis, to get the overall spending overview all at once. To avoid the curse of dimensionality, we propose the innovative use of a DEA-like composite indicator. To be in line with the first stage, we look for a composite indicator that can reflect the idea of the DEA approach: that is, selecting the most advantageous weights for the individual municipality under analysis. Among the composite indicators proposed in the literature, the Benefit-of-the-Doubt model fulfils this requirement, being a data-oriented method that assigns the weights in an objective way to each unit at individual level (Cherchye et al., 2007; Melyn and Moesen, 1991). By construction, the assigned weights' system is the best for the municipality under analysis: as a consequence, comparison and rankings are not on a common system ground. To address this concern and to enable policy makers to evaluate all the municipalities on the same scale with a common system of weights, in the third stage we integrate our model specification with the solution proposed by Despotis et al. (2002) and advocated by Zanella et al. (2013): the composite indicator score of all the municipalities is maximized at the same time, using the same weights across all the units and endogenously obtained within the algorithm specification.

On the other hand, as concerns the effectiveness analysis, econometric models have been proposed to study how specific circumstances or implemented policies might affect the economic phenomenon under investigation. However, the endogeneity issue pervading this kind of analysis prevents a causal interpretation, granting only a correlation one. Several reasons, such as omitted variables, reverse causality, measurement error or sample selection might lead to biased estimation. To unravel a causal relationship, econometric techniques such as Instrumental Variables, Regression Discontinuity Designs, Difference-in-Differences and Fixed Effects, among others are proposed to address this issue (for more details, see Angrist and Pischke, 2008). The listed quasi-experiments rely on several sources of exogeneity, exploiting the geographical location, the legal or political institution setting, the administrative rules as objective thresholds or natural events occurred randomly, for example. Chapter 5 proposes the use of two intertwined policy evaluation techniques, namely a difference-in-differences estimation and Instrumental Variable strategy, to explore the impact of school infrastructure spending on students' achievement. In this case, we rely on the information provided after the earthquake occurred in the Northern part of Italy in 2012. In the educational performance analysis, there are at least two sources of endogeneity (for a more extensive list, we refer to De Witte and López-Torres, 2017): omitted variable and self selection bias. In fact, while evaluating the students' achievement as the outcome of a process using many resources, we should acknowledge that not all the inputs are observable, such as the innate ability of each pupil, the motivations and other family information that is not observable. Moreover, there could be also a problem of self selection if parents might decide where to send their kids to school, for example depending on the school physical environment or on the reputation of the school. The seismic events represent an external shock that enables us to compare similar schools with different entitlement for extra-funding: the exogeneity introduced by the earthquake justifies the assumption of random selection for funding eligibility and allows us to give a causal interpretation to the effect of school infrastructure investment on educational attainment. Specifically, the difference-in-differences approach exploits the information on the allocation process, that is whether schools received funding or not; the Instrumental Variable strategy uses the amount of funding that each school received as function of pre-determined seismic risks.

Before concluding this section, another methodological aspect deserves to be mentioned. In fact, ”considering contextual information makes it possible to understand the major institutional differences and similarities amongst countries, and thereby help to identify comparators for benchmarking purposes” (OECD, 2017c). As unanimously acknowledged, the background conditions matter. Both in the efficiency and the effectiveness analysis, the operating context is addressed in the model specification, in such a way that also this feature can be taken into account. In the efficiency literature, several approaches have been proposed (for an extensive review and critical discussion, see for all Bogetoft and Otto, 2010; Daraio et al., 2017; Simar and Wilson, 2007): among others, non-parametric tests evaluate group differences, they are useful when the underlying distribution is unknown, but may suffer from limited power; bootstrap-based inference makes bias-corrections for the efficiency score computation, but it still assumes the “separability condition”; conditional analysis includes in one-stage the background variables avoiding the assumption of the “separability”, but “if separability holds, the unconditional estimators converge faster than the conditional counterparts” (Daraio et al., 2017). Also in the effectiveness analysis the background variables are directly included in the regression model as control variables, to capture the heterogeneity among the units under observation. In this dissertation we show the possible insights from the application of each of these techniques. In Chapter 2 we estimate the conditional version of the flexible directional distance composite indicator, so that the environmental variables are directly included and the direction of the influence of those variables on the service provision level assessment can be investigated. In Chapter 3 we perform non-parametric significance tests, namely the Mann-Whitney U test to evaluate two group differences, while the Kruskal-Wallis test for three groups or more (see e.g. Kruskal and Wallis, 1952; Ruxton and Beauchamp, 2008, for further details). In Chapter 4 we run a Tobit analysis using bias-corrected efficiency scores. In Chapter 5 we directly include some of the school variables in the model specification as control variables.

It is worth concluding this section by highlighting how different methodologies have been used to answer different research questions. Peculiarities of the institutional context, data availability and policy relevance have boosted the introduction of rather different innovative techniques; notwithstanding, they should be considered in a complementary way so to get a broader view of the public sector management.

1.3 Institutional context

This dissertation presents an empirical application of the proposed enhanced methodologies for two EU countries, Belgium and Italy. These two countries are interesting study cases for their common and specific features, apt to provide complementary insights and informative evidence.

In 2017 both countries provided the Voluntary National Review, available on the Sustainable Development Knowledge Platform, to keep track of the progress, to share their experiences and to foster the 2030 Agenda attainment. Specifically, both Belgium and Italy have a similar interest for the local government analysis and they recognize the importance of the bottom up approach for the Sustainable Development Goals achievement, implementing the global challenges at sub-national level. For example, in Belgium "one in five Flemish municipalities have already signed up to the Global Goals, Local Focus Declaration, thereby acknowledging the importance of the SDGs and the need to develop local actions in support thereof. Pilot projects have been launched in 20 municipalities aiming to fully integrate the SDGs in overall policy and long-term plans by October 2018" (United Nations High Level Political Forum, 2017). [2] Another indicative example is the project called "Health for all" of Tuscany and Tunisia, funded by the United Nations and co-funded by the Tuscany Region: it has been presented at the European Development days in Brussels in the session entitled "Migration, cities and the SDGs. Local authorities key role in implementing the migration-related targets of the Sustainable Development Goals", as an example of the local and regional crucial role in managing the migration and development puzzle.

Furthermore, both Belgium and Italy share a long-standing interest in local government efficiency analysis as witnessed by the related existing literature and specifically in the municipal size issue. An intense wave of mergers occurred among Belgian local governments between 1975 and 1983, restricting the number of municipalities from 2,663 to 589. [3] Then, at the turn of the 21st century a new wave of mergers have been planned and announced in Flanders. [4] Specifically, municipalities have been encouraged to merge on a "voluntary" basis and supported with financial and legal instruments (see for example Sadioglu, 2016) Also in Italy the municipal merger is still a highly debated and ongoing topic. Municipalities are reluctant, especially because the local identity is very strong and deeply rooted in the past: this is the case for Tuscany, for example. [5] The difference not only in terms of municipal size, but also in terms of sub-national expenditure between the two countries is evident from the statistics presented by the OECD (2017d) and reported in Appendix **??**.

In particular, Flanders in Belgium and Tuscany in Italy have more or less the same number of municipalities, but their heterogeneity offers an interesting framework for different and separate analysis, presented re-

[2] See for example https://www.herent.be .

[3] For further details, https://en.wikipedia.org/wiki/Fusion_of_the_Belgian_municipalities

[4] Two Flemish municipalities have very recently decided to merge and others have already announced the merger by 2018 (http://www.flanderstoday.eu/politics/fifth-column-united-we-stand).

[5] For further details, https://it.wikipedia.org/wiki/Fusione_di_comuni_italiani

spectively in Chapter 2 and in Chapter 4, providing new tools of analysis and evidences addressing issues related to the SDG 11 (“Make cities and human settlements inclusive, safe, resilient and sustainable”). For Flemish municipalities the service level provision is estimated by the use of an innovative flexible directional distance composite indicator, identifying the best practices and the factors that mainly affect the overall evaluation. The obtained service provision index is then used to explore its relationship with the local government size and to calculate the optimal municipal tax rate. As regard Tuscan municipalities, the overall level of municipal spending efficiency is computed by means of an original 3-stage non-parametric model and in particular the effect of the municipal size is investigated, given its topical relevance in the policy and academic debate.

Still referring to Tuscan context, the empirical analysis of Chapter 3 involves the wastewater treatment plants controlled by Acque SpA, a public-private utility entrusted in 2002 with water services, and located in the so called “Basso Valdarno” river basin in the Pisa province. This chapter deals with the SDG 6 (“Ensure availability and sustainable management of water and sanitation for all”) and specifically the key role of environmental sustainability in “transforming our world” by 2030. We provide a novel integrated approach to include pollution directly in the efficiency assessment of the plants under analysis: if this aspect is not considered, the “environmentally oriented” plants would be penalized. Moreover, the public-private ownership of the water utility Acque Spa signals an important phenomenon occurring at international level, as for example emphasized in the recommendations promoted by the OECD Council (OECD, 2014): Principle 6 states the need to “mobilise private actors and financing institutions to diversify sources of funding and strengthen sub-national capacities”, so as to bridge the infrastructure financing gap and to develop public-private partnerships (PPP) at the sub-national level, overcoming the infrastructural and service necessities of a growing global population and the limited available resources in an efficient and effective way.

Finally, for Chapter 5 the relevance of education facility upgrade for effective learning environment is considered, referring to the tasks concerning the SDG 4 ("Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all"). For this topic, Italian state high schools are considered, not only because of the particular methodological framework used to evaluate the impact of capital spending on educational achievement, but also because the Italian context offers an interesting setting for other reasons enumerated as follows. First of all, in Italy infrastructure spending in state high schools is mainly a competence at regional level. Additionally, more and more initiatives emphasize the need for more and better school building maintenance and call for increasing capital expenditure in schools: the "XV Report On Safety, Quality And Accessibility To School" provided by a national costumers' organization called "CittadinanzAttiva" has denounced on the news the inadequate maintenance of schoolhouses, reporting that one in four schools lacks of anti-seismic measures and more than ten years would be required to repair the school buildings all over Italy (Corriere della Sera, 28/09/2017). Moreover, in Italy the share in capital expenditure is equal to 2%, way lower compared to 7% as the OECD average and to 8% and 10% in other European countries of similar size, such as respectively France and Germany (OECD, 2017b). To address this need of school infrastructure investment, the Italian government has started financing schools all over the country with a huge amount of money, to keep them safe and adequate, to reinforce the anti-seismic measures and to ameliorate the building system (above all, the "safe schools" program in 2014 and "good schools" law in 2015). However, despite the remarkable pressure to increase capital spending, the effect on the students' outcome is still questioned in the related literature. Last but not least, the schools belong to three different regions, namely Veneto, Emilia-Romagna and Lombardia: these are the regions affected by the seismic events occurred in 2012, for which a large amount of funding has been given to repair and secure safety across the school buildings, giving us the setting for an interesting quasi-experimental design.

1.4 Outline of the dissertation

The dissertation is structured in two main parts and it covers six chapters, as displayed in Table 1.1. **Part I** deals with monitoring basic service provision and measuring its performance level, presenting first an analysis encompassing all the main municipal competencies and then focusing on a specific public service. The research questions of Chapter 2 are:

(I) ***How to measure the overall municipal service level provision in a single index?***

(II) ***Is there an "inverted U-relationship" between production and government size?***

In line with the subsidiarity principle, local governments have experienced an increasing level of involvement in service provision, as closer to the citizens and more suitable to capture their specific needs. The need for more enhanced tools to evaluate the service provision leads us to answer these two questions. We introduce a flexible and fully non-parametric composite indicator, suitable to encompass in one single score multiple municipal intervention areas. While constructing this index, several issues are taken into account, mostly arising from the fact that municipalities differ among each other for a number of reasons. Specifically, regarding the political preferences across the municipal activities, the local budget composition information is directly included in the model specification. Once a synthetic index is obtained as such, we explore a long-standing debate, namely whether there exists an optimal government size, considered here in terms of tax burden imposed on the citizens. To add further insights to the analysis, the influence of few relevant municipal characteristics on the service evaluation is explored, by means of a conditional version of the proposed model specification.

In the next chapter, we focus on a specific service sector, namely water management. In particular, we analyse the performance of the plants in charge of treating wastewater (Wastewater Treatment Plants, WWTPs). The research questions of Chapter 3 are:

(III) ***How to evaluate the wastewater treatment plants performance taking into account environmental sustainability aspects?***

(IV) ***Which are the main determinants affecting the wastewater treatment plants' performance?***

In public sector management, the attention of stakeholders is not only related to performance evaluation in economic terms, but also in terms of sustainability. Accordingly, we first propose a model that combines the benefits of the Non-radial Directional Distance Function and the Analytic Hierarchy Process approaches. As from the first approach, we are able not only to model inputs and good outputs, but also to control for undesirable outputs. Specifically, in our data specification we identify the nitrogen left in the outgoing water as an undesirable pollutant, given its harmful effect on the environment. In addition, from the Analytic Hierarchy Process approach, we include the policy-makers priorities directly in the model specification. In particular, two different scenarios are proposed to show the differences in the evaluation process while taking or not into account to which extent the plants are "environmentally" oriented, so to address the environmental sustainability aspect. Furthermore, several characteristics of the WWTPs are explored, so to give the policy-makers informative suggestions as concerns the main determinants of the WWTP process and their performance.

Part II considers the public spending analysis side, evaluating first the municipal expenditure efficiency across all main local government intervention areas and then assessing the effectiveness of a specific sector funding. The research questions of Chapter 4 are:

(V) ***How to assess municipal expenditure efficiency, considering municipal competencies both one by one and all together?***

(VI) ***How does municipal demographic size affect public spending efficiency?***

In line with New Public Management theories, the efficiency and effectiveness analysis of the public sector is a useful step to support the policy maker decisional process and ultimately to provide better and more tailored services to the citizens. Accordingly, in this chapter we propose a 3-stage DEA based model to evaluate the municipal expenditure efficiency considering the main municipal intervention areas both at non-aggregate and aggregate level. This kind of assessment lends itself well to further investigations. In particular, we address the ongoing debate about the optimal municipal size, while exploring at the same time the impact of other municipal features.

In the next chapter, we consider instead public spending in a specific sector, namely the expenditure in the education sector. In particular, we focus on school infrastructure investment and the way it affects educational achievement. The research questions of Chapter 5 are:

(VII) ***How can we disentangle the endogeneity arising while assessing the school resource effectiveness?***

(VIII) ***Does spending on school physical infrastructure affect student outcomes?***

The need to invest in education and, among others, in school infrastructure is unanimously acknowledged. Nevertheless, despite the evident necessity, the effect of resource provision on educational achievement is still debated. Since the Coleman Report (Coleman, 1966), there is no conclusive evidence as concerns the impact of investment on student outcome, whether supportive or harmful, mostly depending on the endogeneity arising in the standard econometric techniques use. For this reason, we employ a quasi-experimental design and use information on the extra funding that a specific group of schools received in the aftermath of the 2012 Northern Italy earthquake, to investigate the school infrastructure effectiveness with respect to student achievement. Specifically, we use a difference-in-differences strategy to detect whether receiving funding or not affects the educational attainment. An instrumental variable strategy explores the intensity effect depending on the amount of received funding.

Finally, Chapter 6 concludes this dissertation. The main evidence obtained for the raised research questions is summarized emphasizing the policy relevance aspects. To conclude, future lines of research are presented.

Table 1.1: Dissertation outline

	Chapter	Research question	Method	Target groups
	Chapter 1	General introduction		
Part I Service level provision	Chapter 2	How to measure the overall municipal service level provision in a single index?	Robust conditional Directional Distance Benefit-of-the-Doubt with ARI restrictions model	Flemish municipalities
	Chapter 3	How to evaluate the environmental efficiency of wastewater treatment plants including undesirable output?	Integrated Non-Radial Directional Distance function and Analytic Hierarchy Process approach	Tuscan wastewater treatment plants
Part II Public spending analysis	Chapter 4	How to assess the overall efficiency of municipal expenditure and which is its relationship with the municipal size?	Three stage Data Envelopment Analysis based model	Tuscan municipalities
	Chapter 5	Does spending on physical infrastructure affect student outcomes?	Difference-in-differences estimation and instrumental variable strategy	Italian schools
	Chapter 6	General conclusion		

Part I
Service level provision

Chapter 2

Service Level Provision in Municipalities: A Flexible Directional Distance Composite Indicator

2.1 Introduction

In line with the subsidiarity principle and the New Public Management theories, a gradual decentralization of the key activities from the national level to the municipal level has occurred to provide services closer and more tailored to citizens' needs. Accordingly, the pressure on the provision of public goods calls for a more enhanced service level analysis at the local level and, in particular, for suitable tools to measure and monitor local municipal service provision leading towards effective, innovative and sustainable public sector management. Chapter 2 focuses on the evaluation of the overall service provision at municipal level, while Chapter 3 deepens the public goods supply in a specific sector, that is the water management sector.

Specifically, in this chapter we propose an innovative fully non-parametric approach to assess in a dynamic framework the local service provision.

As the municipal tasks cover several multifaceted areas, the proposed method relies on the construction of a composite indicator, and, more specifically, a municipal service provision composite indicator. However, when defining the municipal service provision composite indicator, there are five issues that we need to tackle in the model specification. First, it is necessary to acknowledge that municipalities differ in the activities they develop and do not develop. This decision is often driven by political preferences over the different municipal competencies. This variety is reflected both in terms of local government priorities and in terms of their peculiar specializations: the municipal budget allocation properly keeps track of this kind of information. On the one hand, budget shares are the result of historical choices made by previous governments; on the other hand, they reflect the current government preferences over different municipal intervention areas, depending on municipal characteristics, voter preferences and perceived local needs. Second, the municipalities differ not only in what they are willing to do, but also in what they are able to achieve and the service they can provide. To deal with this kind of variety among the municipalities under assessment and to grant the fairness of comparison to each of them, we propose a flexible approach that grants in a fully non-parametric framework each municipality the benefit of the doubt. The approach is combined with the weight restrictions based on the municipal expenditure composition. In this way we can provide objectively determined and endogenously flexible weights, but at the same time we directly constrain them according to the municipal balance sheets' information. Third, "more" is not always "better" along the dimensions we evaluate the municipal service. For example, the municipalities have the duty to prevent criminality: in this case, the higher the level of criminality, the poorer is the level of service the municipality provides to its citizens in terms of public safety. We deal with this kind of indicators considering them as undesirable features. For this reason, we tailor the suggested BoD approach to a directional distance function as proposed by Zanella et al. (2015). Fourth, the municipal operating environment characteristics play a role in the public activities delivery, as they influence the choice and the importance of the different municipal areas. To avoid the assumption of the "separability condition", we perform a conditional analysis of the emerging model, combined with its robust version, to handle the bias stemming from the atypical observations possibly present in the units under analysis. Finally, as time matters, a dynamic component is added in the conditional model (Mastromarco and Simar, 2015) to exploit intertemporal variations in public service provision (Cordero et al., 2016).

Taking into account the listed issues, we propose an innovative way to evaluate local municipal service provision. More precisely, we advocate a composite indicator built on a directional distance BoD model, including undesirable features and weight restrictions based on the expenditure composition, performing the robust and conditional analysis, within a dynamic framework. The composite indicator is applied to Flemish municipalities to measure municipal service provision of 307 Flemish municipalities over the years 2006–2011.

In a next step we relate the service level provision to local government size. There is significant literature comparing the optimal government size among countries (see for example De Witte and Moesen, 2010), but not among municipalities. Broadly speaking, the main idea is to test the existence of an "inverted U-relationship" between the production (in terms of economic growth or economic performance) and government size, expressed as the share of public sector for the unit under assessment (in terms of % of expenditure, % of revenues or in terms of tax burden). In the economic literature, the theory behind the Armey and the Laffer curves well fits this concept. In the present application, we can consider the tax burden, measured as the municipal tax revenues over the taxable income, against the constructed municipal service composite indicators, to test the existence of a trade-off.

Despite the fact that in the Operational Research literature there is a huge amount of studies focusing on local governments and their efficiency aspects (for an extensive review see Narbón-Perpiñá and De Witte, 2017a,b), the present chapter represents a step forward in several directions. First, there are no studies that measure in a comprehensive way the overall municipal service provision (see for all Fusco et al., 2017;

Karagiannis, 2017). There are studies either proposing this approach only for specific municipal functions, such as for example the waste collection service (Rogge et al., 2017), or focusing only on the global output assessment and not including undesirable features in the evaluation (Afonso and Fernandes, 2008; Yusfany, 2015). Second, when constructing the composite indicator we include the information on the expenditure composition share for each municipal area in the weight restriction specification and not only as a direct weighting scheme to aggregate the municipal tasks (Bosch et al., 2012; D'Inverno et al., 2017; Helland and Sørensen, 2015). Finally, with respect to the huge amount of municipal efficiency papers, just few of them include the robust and conditional analysis (Asatryan and De Witte, 2015; Cordero et al., 2016), even if local public services depend on the characteristics of the municipalities and a fair analysis should account for these differences directly in the main model specification.

The remainder of this chapter is structured as follows. Section 2.2 discusses the different municipal tasks and it explains accordingly the data choice. In Section 2.3 the methodological steps leading to the advocated 'robust conditional directional distance BoD with weight restrictions' composite indicator for the municipal service provision assessment are discussed. Section 2.4 presents the main findings obtained from the empirical application on Flemish municipalities and the usefulness of such proposed composite indicator in exploring the influence of both the municipal background conditions and the government size. Lastly, Section 2.5 presents some final remarks and conclusions.

2.2 Municipal service level

As suggested by the OECD (2008), when constructing a composite indicator we first have to define its theoretical framework, to clarify which is the phenomenon we want to measure and which are the sub-components that can represent it as a whole. Accordingly, we have to make clear which are the dimensions the municipal service composite indicator we propose is built on. The services provided by the municipalities vary from country to country depending on several factors, such as for example the location, the geography, the history and the tradition. However, there are several commonly acknowledged functions that represent the main tasks of a municipality. These include general administration, culture, education and care services, housing and public safety, road maintenance and environmental management. For the sake of clarity, we refer to Figure 2.1 for a list of services grouped by the various functions. These different intervention areas can be seen as the broad categories along which municipal services' composite indicator should be assessed.

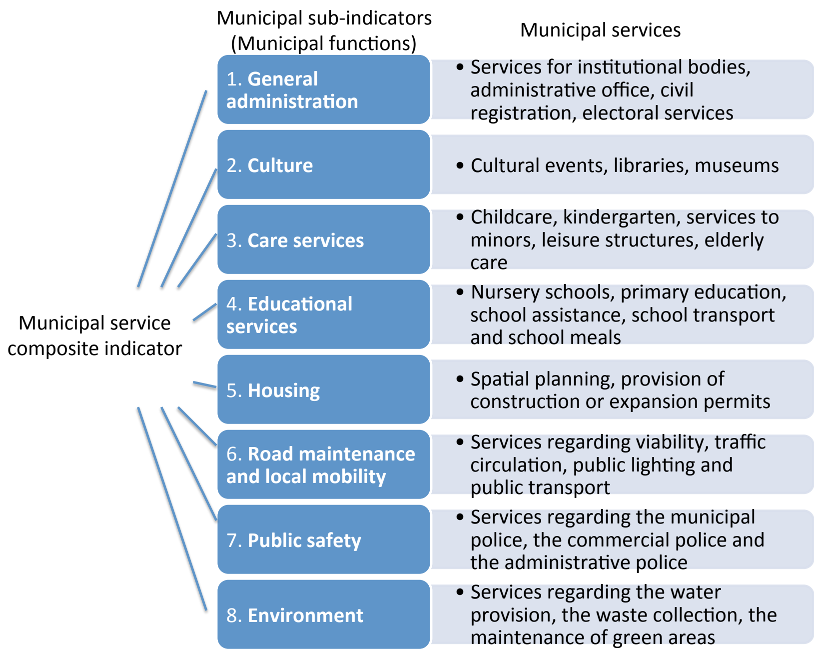


Figure 2.1: Example of services for each municipal function

Once the sub-indicators are identified with the municipal functions, we have to consequently choose suitable and representative variables, in compliance with the data availability and the output choice in the related local governments' efficiency literature (for an extensive review, we refer the interested reader to Narbón-Perpiñá and De Witte, 2017a).

For expository purposes, we focus on the Flemish region of Belgium for which we have exceptionally good data at municipal level. The data refer to 307 Flemish municipalities over the period 2006–2011. Table 2.1 shows the descriptive statistics of the variables expressed in per capita values. We proxy the general administration by the number of *Net Foreigners* and *Households*. The *cultural events* measure the cultural function indicator. The recipient of the education and care services are respectively the *Students in primary school* and the *Children in kindergarten* together with the *Residents over 80*. The *Built-up area* is considered for the housing and country planning area. For the road mobility we consider the number of *Accidents*. For the police function the crime level is measured by the number of *Thefts*, *Physical* and *Property crimes*. Finally, the *Energy consumption* and the *Waste* production are taken into account for the environmental management function. [1]

[1]We selected the listed variables as proxy for municipal functions, but also other indicators can be interestingly used in the composite indicator construction. For example, for Flemish context some indicators used as criteria for funds allocation could be considered (see e.g. http://www.vvsg.be/Werking_Organisatie/Financien/Pages/default.aspx) and used as scope for further research.

Table 2.1: Descriptive statistics for the municipal service level

Municipal service (per 1000 inhabitants)	Obs	Mean	St. Dev.	Min	Max
1. Robust BOD CI for general administration	1842	0.85	0.11	0.73	2.92
Net foreigners	1842	7.44	6.31	0.04	120.8
Households	1842	404.5	23.76	351.33	560.83
2. Cultural events	1842	6.53	5.51	0.3	36.05
3. Robust BOD CI for care service	1842	0.8	0.11	0.56	1.54
Children in kindergarten	1842	37.35	7.37	8.38	96.64
Residents over 80	1842	46.26	9.65	20.19	88.28
4. Students in primary school	1842	63.51	13.47	9.81	145.58
5. Built-up area (Km^2)	1842	64.33	21.2	22.34	251.01
6. Accidents	1842	4.44	1.57	0.39	13.37
7. Robust BOD CI for Crime	1842	0.5	0.21	0.11	1.62
Thefts	1842	21.77	12.42	3.44	92.79
Physical crimes	1842	4.91	2.44	0.9	19.56
Property crimes	1842	8.43	3.56	0	34.01
8. Robust BOD CI for environment services	1842	0.61	0.12	0.41	1.46
Waste ($Tonnes$)	1842	141.08	39.51	56.83	362.63
Energy consumption	1842	7.9	1.8	4.68	22.68

Note: Panel for 307 Flemish municipalities over 2006–2011.

A few additional considerations deserve mention. A recurrent issue in the non-parametric analysis is the curse of dimensionality, occurring when considering a high number of variables in the model. To exploit the data availability and to gather together several aspects of the same phenomenon, a few solutions have been proposed in the literature, as for example aggregating first in sub-indicators (Afonso et al., 2005) or specifying the weight restrictions on a more aggregate level (Morais and Camanho, 2011). In the present application, for some municipal sub-indicators a 'robust Benefit-of-Doubt' composite indicator has been constructed aggregating variables belonging to the same municipal function (for the methodological details, please refer to Section 2.3). This is the case for the following functions: the general administration (composite indicator from Net foreigners and Households), the care services (composite indicator from Children in kindergarten and Residents over 80), the public safety by a crime index (composite indicator from number of

Thefts, Physical and Property crimes [2]), the environmental management (composite indicator from Waste and Energy consumption).[3]

Second, not all the variables included in the analysis are strictly direct measures of the services provided to the citizens, but rather proxies. This procedure is widely accepted both in the composite indicator and in the local governments' efficiency literature to the extent that the selected variables are clearly representative of the intended composite indicator (Narbón-Perpiñá and De Witte, 2017a; OECD, 2008). For example, in the current application the number of kids in the kindergarten and the number of elderly people are used to define the care services function, despite the fact that they are not direct measures of the actual services, as the number of meals per pupil or residential aged care might be for instance. Nevertheless, it is reasonable to assume that the chosen two groups well represent the recipient of this kind of services. Moreover, to the best of our knowledge, there are no available data on the quality of the services and on the citizens' coproduction role, even though this type of information would add a very interesting dimension to the overall analysis, as pointed out for example by De Witte and Geys (2011, 2013).

Finally, it is worth pointing out that not all the municipal sub-indicators represent desirable features, in the sense that "more is better". In fact, the municipalities should not only offer the greatest amount of services they can, but in some cases the service they are supposed to deliver is to contain as much as they can the production of undesirable features, such as the number of accidents, the level of criminality and the level of environmental pollution. To keep the production of these undesirable indicators as low as possible, municipalities have to spend resources that would have otherwise spent in producing other services. For example, municipalities have to pay more subsidies to reduce the level of environmental pollution, in terms of energy consumption and waste production.

[2] Providing a BoD index for crime rather than just adding numbers captures more information: it takes into account the heterogeneity among the different kinds of crime across the municipalities that would be otherwise wiped out.

[3] The robust BoD composite indicator has been computed for every year separately. For the choice of m we choose m=40 and for the bootstrap replications we consider B=2000. Further details on the choice of m and B are provided in Section 2.3.4 and 2.4.

In a similar vein, a local government has to spend resources to keep safe the roads so to minimize the number of accidents and to pay the public officers so to protect and help the citizens. Moreover, despite the fact that quality data are not available as outlined above, the introduced "undesirable" outcomes might reflect to some extent quality (e.g., accidents might partly reflect poor-quality roads). In the current analysis we consider three undesirable and five desirable indicators, the first referring to public safety, road mobility and environmental management functions, the second referring to general administration, education and care services, culture and housing functions.

In addition to the considerations made so far, it is necessary to acknowledge that also the operating conditions matter when constructing a municipal service composite indicator. In fact, the characteristics of the municipalities (e.g., size, income, age composition) affect local public activities and, as a consequence, also the overall assessment of the aggregate indicator. For this reason, in compliance with the variables used in the related literature (for an extensive review, see Narbón-Perpiñá and De Witte, 2017b), three groups of background variables need to be included in the analysis: economic-financial characteristics, socio-demographic structure and the political dimension. We propose the *fiscal income*, the *financial debt* and the *unemployment* as representative variables of economic-financial characteristics: specifically, they are also informative about the institutional setting in which municipalities have to operate. The fiscal income is defined as the income per capita and it represents citizens' economic level estimated for each municipality. Financial debt is measured as the excess of expenditures over revenues per capita and it reflects three interconnected aspects: namely the level of made loans, the return on investment and the fiscal revenue capacity. As stock of deficit, it might be due to decisions rooted in the past, but still it might influence the current service provision. The unemployment is the percentage of unemployed residents between 15 and 64 years over the working population and it can be seen not only as a cost for the municipality in terms of social and housing benefits, but also as a signal of the living conditions in which municipalities have to operate. To frame the socio-demographic structure, we consider the *residents over 65*, the *foreigners* and the *population growth*. The residents over 65 is the share of retired people over the population and it represents the age composition. To capture how a municipality is attractive for foreigners and its ethnic composition, the share of immigrants is considered. Population growth is the variation of residents that a municipality faces over the years and it measures whether the provision of services keeps pace or not with the population growth. Finally, as regards the political aspect, the "*Ideological Complexion of the local Government*" (ICG) measures the ideological stance of the local government on a Left-Right scale (from 0 to 10): a higher ICG score represents a more right-wing government[4]. We refer to Table 2.2 for the descriptive statistics of the presented background variables and Appendix A.1 for additional information on data source and description: some of them are categorized for methodological reasons (for more technical details, see Rogge et al., 2017, and the references therein).

Table 2.2: Descriptive statistics for the municipal background conditions

Background conditions	Obs	Mean	St. Dev.	Min	Max
Economic-financial components					
Fiscal income (€ per capita)	1842	16330.52	1924.97	11055.29	24278.23
Financial debt (€ per capita)	1842	1014.43	581.88	-1497.33	5402.22
Unemployment	1842	5.52	1.8	2.11	15.19
Socio-demographic components					
Residents over 65 (% of total)	1842	0.18	0.02	0.11	0.3
Foreigners (% of total)	1842	0.05	0.06	0	0.48
Population growth	1842	0.64	0.62	-4.51	3.59
Political component					
Ideological Complexion of the local Government (ICG)	1842	5.04	0.71	2.5	6.3

Note: Panel for 307 Flemish municipalities over 2006–2011.

2.3 Methodology

This chapter proposes an enhanced way to measure the local municipal service provision by constructing a composite indicator. As there

[4]We gratefully thank De Witte and Geys (2009) for providing us this data.

are several issues that we need to tackle while specifying the model, the methodological steps leading to the advocated 'robust conditional directional distance BoD model with weight restrictions' are presented in the next subsections.

2.3.1 The BoD model

Local municipal service provision covers several areas, as presented in the previous section. Accordingly, we look for an aggregating method to group several dimensions into one single composite indicator (CI). As we do not have a priori knowledge of the functional form and understanding of the importance of the different municipal services, we consider a fully non-parametric way to avoid any kind of specification bias. In particular, we choose a Benefit-of-the-Doubt (BoD) weighting technique, inspired by the Data Envelopment Analysis (DEA) methodology (Charnes et al., 1978) and labelled as such after Melyn and Moesen (1991). The peculiarity of the BoD model is that it assigns endogenously the weights for each municipal service. More specifically, the service provision level of the municipality under analysis is compared in a relative perspective to the service level of all the municipalities in the sample: a higher weight is assigned to a municipal area where the municipality under analysis provides relatively high service level and a lower weight where it provides relatively low service level. To put it differently, the set of weights is defined so that high importance is assigned to municipal areas with relatively high level of services and low importance is assigned with relatively low level of services. The optimal weights are determined in such a way that the composite indicator for the overall level of service provision of the municipality j_0 under analysis is maximized and they are obtained solving for each municipality j_0 the following problem:

$$
\begin{array}{lll}
CI_{j_0} = \max & \sum_{r=1}^{s} y_{rj_0} w_{rj_0} & \\
\text{s.t.} & \sum_{r=1}^{s} y_{rj} w_{rj_0} \leq 1, & \text{for } j = 1, \ldots, j_0, \ldots, n \\
& w_{rj_0} \geq 0, & \text{for } r = 1, \ldots, s
\end{array}
\tag{2.1}
$$

with CI_{j_0} the composite indicator optimal value for the evaluated municipality j_0; y_{rj_0} denotes the observed service level for the municipal area r of the evaluated municipality j_0; w_{rj_0} the most favorable weight to the municipal area r for the evaluated municipality j_0; y_{rj} the observed service level for the municipal area r of every municipality j in the dataset; n the number of municipalities under analysis (n=307) and s the number of municipal functions considered in this application (s=8).

The first constraint in the model formulation is referred to as the "normalization" constraint: the overall municipal composite indicator CIj_0 is maximized subject to an upper bound equal to one. Therefore, the CIj_0 value ranges between zero and one: the higher the value, the higher is the overall service provision level for the evaluated municipality. If CIj_0 <1, it means that, even if the municipality under analysis is evaluated with its most favorable weighting system, there is at least another municipality providing a higher overall level of service. Hence, there is still room for improvement in the service provision, given the observed overall level of provided services across the whole sample. If CIj_0 =1, the municipality under analysis is not outperformed in terms of the overall service provision and it is considered as its own benchmark while using its most favorable weight system. The second constraint imposes the weights' non-negativity.

The advantage of using this approach is twofold: first, it allows to group together several aspects into one single indicator. Second, it ensures the fairness of the comparison, weighting more the municipal areas where higher priority is devoted and, vice versa, weighting less the ones with lower priority. In this way, each evaluated municipality is granted the "Benefit-of-the-Doubt" in the assessment and the fairness of the comparison is ensured (for more details on the BoD approach, see e.g. Cherchye et al., 2007; Rogge et al., 2017; Verschelde and Rogge, 2012).

2.3.2 The directional distance BoD model

In the depicted BoD framework, a higher indicator level in a certain municipal area contributes to a better overall service provision assessment: to this extent, the indicator can be labelled as "desirable". However, we have to acknowledge that this is not necessarily always the case among all the local services. In fact, municipalities might also provide services in areas where the best they can do is to contain the production of the indicator rather than to expand it and for this reason the label "undesirable" is assigned. For example, among the environmental services, municipalities are supposed to promote activities to reduce the waste production and the energy consumption. As a consequence, the municipal waste and the energy consumption enter in the model as undesirable indicators: municipalities have to pay more in terms of subsidies to promote a lower consumption of energy and lower production of waste, devoting resources that would have otherwise spent on other services.

The inclusion of undesirable features in the construction of composite indicators is quite recent and it is linked to the performance measurement literature (for an extensive review, see Dakpo et al., 2016; Zanella et al., 2015). In this study, we propose the model introduced by Zanella et al. (2015) and advocated by Rogge et al. (2017), namely a directional distance BoD model. This model combines the earlier listed advantages of the BoD approach together with the ones of the directional distance function, introduced by Chung et al. (1997). In fact, the directional distance model allows to simultaneously contract the undesirable indicators and expand the desirable ones along a specified direction vector $g = (-g_b, g_y)$, as shown in its primal formulation (Zanella et al., 2015, model (7), p.523). However, the multiplier formulation of the directional distance BoD model (Zanella et al., 2015, model (8), p.523) is preferred to include weight restrictions in the municipal service level assessment and it has to be solved for each j_0 municipality under analysis:

$$
\begin{aligned}
\beta_{j_0} = \quad \min \quad & -\textstyle\sum_{r=1}^{s} y_{rj_0} u_{rj_0} + \sum_{k=1}^{l} b_{kj_0} p_{kj_0} + v_{j_0} \\
\text{s.t.} \quad & \textstyle\sum_{r=1}^{s} g_y u_{rj_0} + \sum_{k=1}^{l} g_b p_{kj_0} = 1 \\
& -\textstyle\sum_{r=1}^{s} y_{rj} u_{rj_0} + \sum_{k=1}^{l} b_{kj} p_{kj_0} + v_{j_0} \geq 0 \quad && \text{for } j = 1, \ldots, j_0, \ldots, n \\
& u_{rj_0} \geq 0 && \text{for } r = 1, \ldots, s \\
& p_{kj_0} \geq 0 && \text{for } k = 1, \ldots, l \\
& v_{j_0} \in \Re
\end{aligned}
\tag{2.2}
$$

with β_{j_0} the optimal value for the evaluated municipality j_0; y_{rj_0} and b_{kj_0} respectively the observed r desirable and k undesirable indicator of the evaluated municipality j_0; u_{rj_0} and p_{kj_0} respectively the most favorable BoD-weights for the r desirable and k undesirable indicator for the evaluated municipality j_0; y_{rj} and b_{kj} respectively the r desirable and k undesirable indicator of every municipality j in the dataset; n the number of municipalities under analysis (n=307); s and l respectively the number of municipal functions linked to desirable and undesirable indicators considered in this application (s=5 and l =3).

Among several direction vectors that can be specified, we choose $g = (-b_{kj_0}, y_{rj_0})$: in other words, we use the municipal service indicators of the evaluated municipality as the direction vector. In this way, the composite indicator for the municipal service provision is obtained as

$$CI_{j_0} = 1/(1 + \beta_{j_0})$$

and it ranges between zero and one, where one denotes the greatest level of service provision as in the basic BoD model.

2.3.3 The directional distance BoD model including weight restrictions

In the local service provision assessment, there is another aspect that can not be ignored, namely the political preferences over the different municipal intervention areas. There are two interconnected explanations for this kind of heterogeneity among the municipalities. First of all, there are municipal functions that deserve higher priorities than others. This phenomenon is not only quite evident looking at the average expenditure composition across the municipalities, but it is also clearly stated in certain national legislative systems (for example in Italy there is the distinction between "fundamental" and "non-fundamental" functions). Second, every municipality has its own peculiar vocation: a municipality might be more focused on the tourism sector, another one on cultural activities, another one on a different economic specialization, just to provide few examples. In this case, the budget allocation reflects variety across the municipalities under evaluation, as they consider as more important.

By including weight restrictions in our model formulation, we can not only include these value judgements, but we can also address one common concern related to the great flexibility in the weighting system associated to the BoD approach. In the DEA/BoD literature several types of weight restrictions have been considered (for a review, see for all Cherchye et al., 2007; Sarrico and Dyson, 2004; Zanella et al., 2015, and references therein). In this context, we suggest the assurance region type I (ARI) weight restrictions as suggested by Zanella et al. (2015) and advocated by Calabria et al. (2016).[5] By adding this kind of restrictions to

[5]As suggested by Sarrico and Dyson (2004), this kind of choice might penalize the units

the model specification, we can constrain the relative importance of each municipal function indicator within a certain range and express it in percentage terms, as follows:

$$\begin{cases} \phi_r \leq \dfrac{u_{rj_0}\bar{y}_r}{\sum_{r=1}^{s} u_{rj_0}\bar{y}_r + \sum_{k=1}^{l} p_{kj_0}\bar{b}_k} \leq \psi_r & \text{for } r = 1, \ldots, s \\ \phi_k \leq \dfrac{p_{kj_0}\bar{b}_k}{\sum_{r=1}^{s} u_{rj_0}\bar{y}_r + \sum_{k=1}^{l} p_{kj_0}\bar{b}_k} \leq \psi_k & \text{for } k = 1, \ldots, l \end{cases} \tag{2.3}$$

with s and l constraints respectively for each observed r desirable and k undesirable municipal indicator.

The question remains on how to specify the importance of each municipal indicator and accordingly the bounds ϕ and ψ. The innovative way we propose in this chapter is to get this kind of information directly from the municipal expenditure allocation across the different services. To the best of our knowledge, in municipal performance assessment the expenditure composition has been included directly in the aggregation process, but not in the weight restrictions (see for example Bosch et al., 2012; D'Inverno et al., 2017; Helland and Sørensen, 2015). In this regard, the expenditure composition of each municipality is consistent with the rationale of the "budget allocation" approach as described by Cherchye et al. (2007).

The proposed method has the advantage to reflect the heterogeneity across municipalities, granting some leeway but, at the same time, leaving an objective order of importance among the municipal services without imposing any kind of external judgement. In particular, we propose three sets of restrictions, which vary according to the different specified bounds.

The first one considers the minimum and the maximum share of expenditure in each municipal area across all the municipalities ("*MinMax* restrictions"). In this way, the municipality under evaluation cannot assign lower or greater importance to each municipal indicator than the one recognized among all the municipalities.

In a second alternative way to specify the restrictions, for each municipal indicator the average spending share is considered, identifying a lower and an upper bound value equal to its $\pm 50\%$ ("*Average* restrictions"). This kind of restrictions circumscribes the average importance

with small or large values. However, as emphasized by Zanella et al. (2015, p. 526), among other weight restriction alternatives the ARI type is "the best option to construct composite indicators and ranks", so to ensure a fair comparison among the units under evaluation.

of each municipal area according to the priorities acknowledged among all the municipalities: local governments are given some leeway in deciding their own weights, but at the same time a certain order of importance among the functions is respected.

Finally, rather than confining a municipality within the overall average choice, the third specification of restrictions allows each municipality to set its own weight based on its current spending allocation ("*Municipal-specific* restrictions"). To put it differently, the lower and the upper bound value of the constrains associated to each municipal indicator is equal to $\pm 50\%$ of each municipal-specific expenditure share.

By construction, the three sets exhibit increasingly binding restrictions. Table A.4 presents summary information about the weights just presented for each municipal area: we refer to Appendix A.2 for the municipal-specific lower and upper bounds.

Table 2.3: Summary of the weights obtained from the municipal expenditure composition

	Administration	Culture	Care services	Education	Housing	Local mobility	Security	Environment
MIN	0.08	0.02	0.01	0	0.01	0.03	0.04	0.03
LOWER BOUND	0.09	0.07	0.06	0.05	0.03	0.07	0.05	0.05
AVERAGE	0.18	0.14	0.12	0.10	0.05	0.15	0.11	0.10
UPPER BOUND	0.27	0.21	0.17	0.14	0.08	0.22	0.16	0.15
MAX	0.49	0.31	0.23	0.39	0.16	0.33	0.23	0.22

Interestingly, the optimally chosen weights do not only reflect the importance that each municipal indicator has in the overall assessment of service level provision, but they can also be interpreted as normalized shadow prices (Coelli et al., 2005). In our application, a shadow price can describe the way a municipal indicator is affected whenever another indicator varies, or alternatively how the composition of overall service provision can change depending on different political choices (for more technical details, see Fusco, 2015; Grupp and Schubert, 2010). The shadow prices are useful to determine the "budget shares" (Van Puyenbroeck and Rogge, 2017).

2.3.4 The robust and conditional directional distance BoD model

All the steps discussed so far are necessary to grant an increasing level of fairness in the local service provision assessment. However, there is a last aspect that can not be neglected, namely the role of the operating context under which the municipalities have to operate. First of all, the background conditions can affect both the supply and the demand side of the service provision level. For example, as concerns the supply aspect, a wealthier municipality with a higher level of local revenues might be endowed with more resources to spend. Alternatively, as for the demand side, a municipality experiencing a higher level of unemployment might be required to provide a higher number of subsidies, diverting resources from the provision of additional services. Moreover, background conditions can also have a remarkable impact on the political preferences over municipal functions, influencing to different extents the components of the composite indicator and the way they enter in the synthetic index. As a consequence, the 'directional distance BoD with ARI restrictions' composite indicator CI_{j_0} outlined so far has to be adjusted accounting for the differences in the municipal environmental variables. These can be grouped in economic-financial characteristics, socio-demographic structure and political dimension.

Despite the fact that the operating factors are exogenous with respect to the service provision level and they are not under the control of local policy-makers, they affect not only the distribution of the composite indicator scores, but also their attainable set. For these reasons, the "separability condition" can not be assumed and we need to use a one-stage procedure to compute the municipal service provision composite indicator including at the same time the environmental factors: in the literature this approach is referred to as the "conditional" measurement procedure. Moreover, we complement the conditional analysis with its robust version to mitigate the influence of outlying observations, arising from, e.g., measurement errors and atypical observations, using the insights from the "order-*m*" approach. For the sake of brevity, we refer for a more formal and extensive explanation of the procedures to Cazals et al. (2002); Cordero et al. (2017); Daraio and Simar (2005, 2007); De Witte and Kortelainen (2013), among others.

For the computation of the robust municipal service composite indicator, we run a Monte-Carlo algorithm performing B computation rounds (where B is large): the bootstrap replicates reduce the impact of the outlying observations. In each b round ($b = 1, \ldots, B$), first m municipalities are drawn with replacement from the original sample of n units and then the m-sample 'directional distance BoD with ARI restrictions' composite indicator $CI_{j_0}^{b,m}$ is computed. Finally, the robust composite indicator $CI_{j_0}^{m}$

is obtained as the arithmetic average of the B $CI_{j_0}^{b,m}$, as follows:

$$CI_{j_0}^{m} = \frac{1}{B}\sum_{b=1}^{B} CI_{j_0}^{b,m}$$

The $CI_{j_0}^{m}$ score might be larger than one: this means that the municipality j_0 under evaluation is providing a higher service level than the average *m* municipalities it has been compared with as its reference sample.

To include also heterogeneity among the municipalities captured by the *z* background variables, the Monte-Carlo simulation procedure is the same but in the drawing. The *m* municipalities are drawn with replacement and with a particular probability based on an estimated kernel density function. The idea is to draw *m* municipalities with a higher probability of being similar to the municipality j_0 under evaluation (and lower probability of being dissimilar): in this way, the municipal service provision level composite indicator

$$CI_{j_0}^{m,z} = \frac{1}{B}\sum_{b=1}^{B} CI_{j_0}^{b,m,z}$$

is assessed considering similar background conditions and ensuring a greater level of fairness in the comparison among the municipalities under analysis. In this case, a $CI_{j_0}^{m,z}$ score larger than one means that the municipality j_0 under evaluation is providing a higher service level than the average *m* municipalities with similar background characteristics, while $CI_{j_0}^{m,z} = 1$ denotes a similar service provision level.

Finally, as we want to extend the model to a dynamic framework to exploit intertemporal variations in municipal service provision (Cordero et al., 2016), we further adjust the robust conditional composite indicator $CI_{j_0}^{m,z}$ including the time dimension, according to the insights of the approach proposed by Mastromarco and Simar (2015). Accordingly, the composite indicator $CI_{j}^{m,z,t}$ is computed over all the combinations of municipality $j = 1\ldots, j_0, \ldots, n$ and time period $t = 1\ldots T$ (where $T = 6$ in our application) and time is also included as an additional conditioning variable together with z.

Furthermore, the computation of the robust unconditional and conditional composite indicators provides two more additional useful insights. By means of a non-parametric statistical inference, we can detect whether the environmental variables are on average statistically significant with respect to the composite indicator scores and which is the direction of the influence of the environmental variables on the service provision level assessment. More specifically, the ratio of the unconditional CI_j^m and conditional $CI_j^{m,z,t}$ composite indicator scores can be non-parametrically regressed on the external variables. By construction, the unconditional CI score is at most (less than or equal to) the conditional CI score. Accordingly, a positive coefficient denotes a favourable effect of a contextual variable on the service provision level score: when the variable increases, the conditional CI score gets closer to the unconditional one. For a negative coefficient, the opposite holds: the contextual variable has an unfavourable effect on the service provision level assessment, as the conditional CI score increases when the variable increases. To put it in an alternative way, the background condition acts as an unfavourable context if, as the variable increases, its CI score increases only because evaluated among similar municipalities and the service provision level it can afford is lower compared with the one of municipalities facing different background conditions. For the sake of brevity, we omit further technical and theoretical details: we refer the interested reader to Bădin et al. (2012); Daraio and Simar (2007).

To recap, in this methodological section the steps leading to the advocated 'robust conditional directional distance BoD with ARI restrictions' composite indicator $CI_j^{m,z,t}$ for the municipal service provision level assessment are presented. To facilitate further research and to allow for practical policy implementation, the code is available upon request.

2.4 Results

In this section we present the empirical application of the proposed method on a sample of 307 Flemish municipalities over the years 2006–2011. First, we make some comments on the evidence stemming from the inclusion of different weight restriction specifications and from the comparison with the robust and conditional versions, including the analysis of the influence of the municipal operating context. Second, we include some findings exploring the relationship between the obtained municipal service composite indicator and the local government size.

2.4.1 The municipal service provision composite indicator results

We show the results of the estimated robust conditional municipal service composite indicator, for different weight restriction specifications (*MinMax*, *Average* and *Municipal-specific*) as presented in section 2.3.3. We estimate different conditional models, depending on the group of background variables as introduced in section 2.2. Model 1 includes the economic and financial characteristics that might affect the municipal service delivery, namely the level of fiscal income, the level of financial debt and the unemployment rate. In Model 2, the socio-demographic structure is also added, by including the share of elderly people, the share of foreigners and the municipal population growth. The political component is considered together with the economic and socio-demographic characteristics in Model 3, by using the Ideological Complexion of the local Government (ICG). Moreover, in every model specification a year dummy is also included to run the analysis in a dynamic framework. For the sake of comparison, the unrestricted unconditional, the unconditional and the robust unconditional models are also estimated. Table 2.4 shows the descriptive statistics of the estimated composite indicator results.

The results can be explored along two complementary dimensions.[6] The first one is related to the use of the weight restrictions. Not surprisingly, the inclusion of the weight restrictions lowers the values of the composite indicators with respect to the unrestricted model: every municipality under analysis is forced to choose its own optimal system of weights only within a certain range. Moreover, for each model specification the three different sets of weight restrictions lead to a lower average service provision: as pointed in section 2.3.3, they are by construction increasingly binding. Including the information on the expenditure composition does play a role in the composite indicator estimation through alternative weight specifications. In addition, further information can be retrieved from the shadow prices, as they are useful to determine the "budget shares" (Van Puyenbroeck and Rogge, 2017). In particular, we can observe that imposing weight restrictions gives back more reliable results as closer with the current composition. Budget allocation cannot be changed drastically, as otherwise suggested by the unrestricted results, and a minimum expenditure share is granted to each function

[6]Results and main inference analysis are obtained by using Matlab 16a, R and Stata 13. The codes are available upon request.

Table 2.4: Descriptive statistics of the service provision composite indicator scores estimated for 307 municipalities over 2006–2011

	Mean	St. Dev.	Min	Max
Unrestricted Unconditional	0.8388	0.0643	0.6357	1.0000
Unconditional				
MinMax restrictions	0.7832	0.0670	0.6302	1.0000
Average restrictions	0.7178	0.0649	0.5883	1.0000
Municipal-specific restrictions	0.7067	0.0723	0.5121	1.0000
Robust Unconditional				
MinMax restrictions	0.9618	0.1097	0.8024	2.0825
Average restrictions	0.8752	0.0937	0.6994	1.5803
Municipal-specific restrictions	0.8663	0.1042	0.6650	1.8782
Robust Conditional Model 1				
MinMax restrictions	0.9753	0.0315	0.8047	1.0012
Average restrictions	0.9215	0.0612	0.7029	1.0008
Municipal average restrictions	0.9158	0.0673	0.6585	1.0001
Robust Conditional Model 2				
MinMax restrictions	0.9969	0.0093	0.8895	1.0000
Average restrictions	0.9846	0.0275	0.7766	1.0000
Municipal average restrictions	0.9832	0.0312	0.7140	1.0000
Robust Conditional Model 3				
MinMax restrictions	0.9983	0.0059	0.9097	1.0000
Average restrictions	0.9906	0.0197	0.7775	1.0000
Municipal average restrictions	0.9893	0.0234	0.7172	1.0000

Note: Unrestricted indicates the absence of weight restrictions. *MinMax* restrictions refer to the minimum and maximum share of expenditure in each municipal area across all the municipalities. *Average* restrictions consider the average spending share (lower and upper bound equal to its $\pm 50\%$). *Municipal-specific* restrictions are based on the municipal-specific current spending allocation (lower and upper bound equal to the $\pm 50\%$ of each municipal spending share).

Model 1 includes the economic and financial characteristics (Fiscal income, Financial debt and Unemployment). *Model 2* adds to the economic and financial characteristics the socio-demographic structure (Share of elderly people, Share of foreigners and Population growth). *Model 3* adds to the economic-financial and socio-demographic variables the political component (Ideological Complexion of the local Government). In every conditional model specification a year dummy is also included.

in this way, avoiding zero weights whenever a minimum amount of investment is required in a certain municipal function. For the full list of results, we refer to Appendix A.4 for the shadow price and the budget share results.

The second dimension refers to the inclusion of robust and conditional analysis.[7] The scores tend to increase when performing the robust and the conditional analysis compared to the unconditional estimation. This evidence shows the methodological importance of such an integrated analysis: the atypical observations and the background variables do affect the composite indicator. Specifically, in the present application when considering the operating environment each municipality has to work in, it turns out that there is no longer much room for improvement left as the mean scores are almost equal to one in the conditional model results.[8] On the one hand the evidence shows that this set of variables explains largely the service level provision in municipalities; on the other hand, the conditional analysis helps in identifying the correlation between some municipal characteristics and the service provision level.

The influence of the contextual variables on the municipal service provision level can be detected by looking at the robust unconditional and conditional estimates together, as explained in section 2.3.4. Table 2.5 presents the results of the statistical inference. For the sake of brevity, only the results for the MinMax restrictions are presented along the chapter: the results are robust across the three weight restriction specifications and we refer to Appendix for the complete list of results. The direction of the influence of the environmental variables is in line with the main evidence described in the literature on local government's efficiency (see for all Narbón-Perpiñá and De Witte, 2017b). Concerning the economic and financial characteristics, the level of fiscal income plays an unfavourable role in the municipal service provision assessment. When local governments have a greater amount of financial resources, the politicians might spend in a less prudent way and the citizens might be less motivated to monitor the expenditures: as a result, the overall level of delivered services seems to reduce. The financial debt has an unfavourable correlation

[7] After a sensitivity analysis for the choice of m (m=10, 20, ..., 100), we choose m=40 for which there is a remarkable decrease of the super-efficient municipalities. As for the bootstrap replications, we consider B=2000.

[8] In Appendix A.3 the Spearman's rank correlation coefficients across the different weight restriction specifications and model distributions are reported to further investigate the impact on the ranking and on the best practices among the municipalities under evaluation.

too. When the level of local government debt is higher, more resources will be spent on debt interests and amortization payments: therefore less resources will be available and this will bring to an overall lower level of service provision. Less amount of resources is available to provide municipal services also when there is a higher level of unemployment: higher spending is devoted to social and housing benefits. Hence, in the overall service provision assessment also this variable plays an unfavourable role.

The dataset covers the 2006-2011 period, which mostly coincides with the term of local authorities elected in 2006. As in Cordero et al. (2016), we adopt a dynamic approach to exploit intertemporal variations in public service provision and to observe whether municipalities made some changes during their electoral term. Interestingly, if we consider the time trend looking at the partial plot for the year variable (see Figure 2.2) combined with the economic characteristics, we can see that 2008, the year of the economic crisis, is the most unfavourable as concerns the municipal service provision (the same is observed in each weight restriction specification). However, from 2009 increases in public service provision have been recorded. This phenomenon might be linked to the fact that in 2009 a new legislative era began at national level. One of the main priorities of the government was to stimulate public service provision at local level in line with the subsidiarity principle (Sadioglu, 2016). Therefore, the negative impact of the crisis might have been balanced by the renewed attention on local service provision.

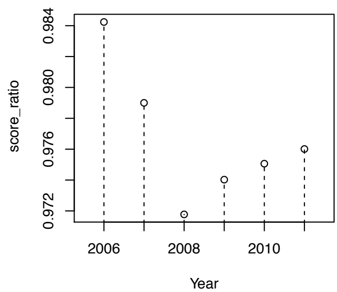


Figure 2.2: Intertemporal variation of the service provision when considering economic variables

When including also the socio-demographic characteristics, the interactions among the variables lead to contradicting evidence related to the direction of the influence of the level of fiscal income and financial debt, playing as counteracting factors. For example, this might be due to the fact that municipalities are more concerned with the public resources management when they have a higher level of accountability or when they pay more attention on cost saving due to their financial problems. In addition, the results of Model 2 on the socio-demographic structure show that the share of elderly people and of foreigners has a favourable effect on the municipal service provision assessment, whenever it is statistically significant. These population groups are the recipients of several municipal services provided with the aim of satisfying their needs: the higher the number of people in these categories, the greater the level of scale economies exploitation. Vice versa, when the level of population growth is too high, the municipalities might not be able to completely satisfy overall citizens' demand and therefore, all else equal, it will lead to a lower level of provided services.

Finally, Model 3 includes also the information on the political component, namely the Ideological Complexion of the local Government (ICG), that captures the ideological stance of the municipality on a Left-Right scale. We observe that a low level of municipal service provision is associated with a more right-wing government. In fact, as a common hypothesis a more left-wing coalition is more prone to have a larger public sector.

Table 2.5: Influence of background conditions on municipal service composite indicator

MinMax weight restrictions

	Model 1			Model 2			Model 3		
	Influence	p-value		Influence	p-value		Influence	p-value	
Economic-financial									
Fiscal income	Unfavourable	0.000	***	Unfavourable	0.000	***	Unfavourable	0.000	***
Financial debt	Unfavourable	0.000	***	Favourable	0.000	***	Favourable	0.000	***
Unemployment	Unfavourable	0.000	***	Unfavourable	0.170		Unfavourable	0.085	*
Socio-demographic									
Residents over 65				Favourable	0.075	*	Favourable	0.000	***
Foreigners				Favourable	0.045	**	Favourable	0.080	*
Population growth				Unfavourable	0.000	***	Unfavourable	0.000	***
Political									
ICG							Unfavourable	0.000	***

Note: The background variable has an *unfavourable influence* on the service provision assessment when the municipal composite indicator score increases only because the municipality under assessment is evaluated among similar municipalities: the service provision it can afford is lower compared with the one of municipalities facing a different context. The opposite holds when a background variable is found to have a *favourable influence*.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

2.4.2 Municipal service provision and government size

Once the overall level of service provision is estimated, we can use the obtained composite indicator score not only for benchmarking and for detecting the best practices, but also to explore the relationship between the service level and several municipal aspects. Among those, referring to the second question raised in the introduction, we want to investigate whether there is a trade-off between the service level and the local government size. For this application, the tax rate is defined as the ratio between "Total tax revenue per capita" (*Totale belastingontvangsten per inwoner*) and "Taxable income per capita" (*Belastbaar inkomen per inwoner*). To test the existence of such a relationship, the regression model to be estimated is the following:

$$CI_j = \beta_0 + \beta_1 GOVsize_j + \beta_2 GOVsize_j^2 + \epsilon_j$$

As evident from the plots in Figure 2.3, the "inverted U-relationship" is recurrent in every model and weight restrictions specification, both for the unconditional and the conditional estimates that take into account the municipal background variables in one-stage: β_2 is almost always negative and statistically significant.

To calculate the optimal municipal tax rate that can maximize the service provision level, the following equation should be used:

$$\frac{\partial CI}{\partial GOVsize} = 0$$

and then the optimal point is computed as follows:

$$GOVsize^* = -\frac{\beta_1}{2\beta_2}$$

Table 2.6 presents the optimal tax rate computed across the different model specifications. The optimal value ranges between 2.04% and 5.30%. The actual average tax rate in Flanders between 2006 and 2011 is equal to 3.75% and it has been computed as the average across 307 municipalities over the years under analysis: it results quite below the optimal average 4.30% computed across the different model specifications. More specifically, when considering both the economic variables and the expenditure structure by using the weight restrictions, the average optimal size is systematically higher than the one set by Flemish local governments over the years 2006–2011. In Appendix A.6, the results are presented also for the conditional model specification encompassing also Socio-demographic and political variables.

Table 2.6: Optimal tax rates across different model specifications

	Unconditional	Robust	Model 1
MinMax	2.04%	3.65%	5.30%
Average	5.00%	4.83%	5.29%
Municipal average	4.07%	3.58%	4.96%

Note: MinMax restrictions refer to the minimum and maximum share of expenditure in each municipal area across all the municipalities. *Average* restrictions consider the average spending share (lower and upper bound equal to its ±50%). *Municipal-specific* restrictions are based on the municipal-specific current spending allocation (lower and upper bound equal to the ±50% of each municipal spending share).

Model 1 includes the economic and financial characteristics (Fiscal income, Financial debt and Unemployment).

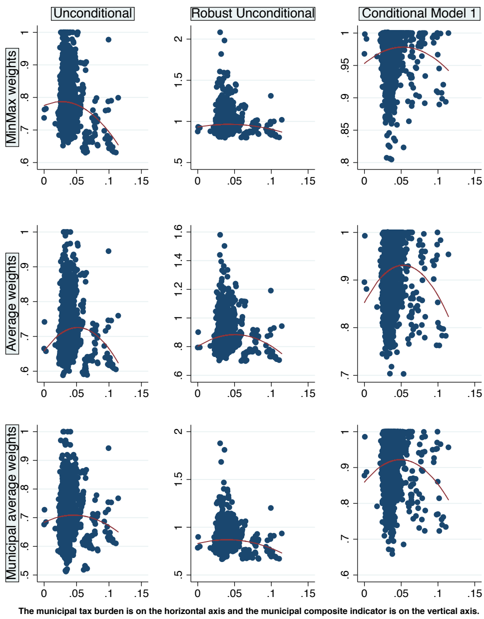


Figure 2.3: "Inverted U-relationship" between municipal tax burden and municipal service provision level

2.4.3 Robustness checks

To conclude this section, it is worth to point out that as a robustness check a further analysis is performed excluding from the sample the largest cities (so-called "Centrumsteden"[9]) to avoid possible criticism on the municipal size effect, even if the robust analysis should have taken this issue already into account. The main findings are confirmed and the results are presented in Appendix A.7.

2.5 Conclusion

In this chapter we propose an innovative tool to measure municipal service provision level, a robust conditional directional distance Benefit-of-the-Doubt with weight restrictions composite indicator. To show the usefulness of the proposed approach, we compute the municipal service provision composite indicator for 307 Flemish municipalities over the year 2006–2011.

The advocated model specification is fully flexible, able to capture the multifaceted aspects involved in local public goods provision evaluation and, in particular, heterogeneity among the municipalities in their activities, to grant a fair analysis. Accordingly, local political preferences and municipal characteristics are directly embedded in the model. Overall, the approach ensures an objective way to determine how each municipal area enters in the evaluation, while granting the most favourable aggregating scheme for the units under analysis. The information on the municipal expenditure composition is included through the weight restrictions specification. The directional distance function formulation makes possible the evaluation even along undesirable features, which should be reduced rather than maximized in the problem. The robust conditional version of the model controls for the municipal operating context and the time dimension.

The proposed composite indicator not only groups together all these components, going a step forward in the existing literature, but it also allows further investigation. First of all, we can explore how municipal characteristics influence overall service provision through statistical inference, detecting whether the background condition inclusion favours

[9]The "Centrumsteden" indicates 13 Flemish city centres, with relatively high numbers of inhabitants, that play a central role in the employment, care, education, culture and recreational activities. `https://nl.wikipedia.org/wiki/Centrumstad`

or not the assessment. More broadly, the obtained composite indicator can be used to explore the relationship between the provided municipal services and some relevant issues in municipal management, as for example government size expressed in terms of the tax burden imposed on citizens, which pay the taxes for the local public goods they receive.

In the empirical application, we find an unfavourable influence in the municipal service provision assessment regarding the considered economic variables, namely the level of fiscal income, financial debt and unemployment. As concerns the share of elderly people and foreigners, a favourable influence is found, while for the population growth the opposite holds. Finally, a left-wing government favours municipal activities. As regards the relationship between the municipal service provision composite indicator scores and the tax burden, an "inverted U-shape" is recurrent in every model and weight restriction specification, confirming the hypothesis proposed in the economic literature: beyond an optimal point, despite a higher level of revenues, a lower level of public goods is provided to the citizens.

The analysis proposed in this chapter has the benefit of framing a broad view of all the services received by the citizens. However, it does not contain information about the quality of the provided services, despite its relevance in the overall service assessment and in terms of citizens' satisfaction evaluation. The next chapter enables us to go deeper in this topic, while focusing on a specific sector though, namely the water sector. In particular, it considers the environmental sustainability in the water management performance measurement.

Chapter 3

Water pollution in wastewater treatment plants: an efficiency analysis with undesirable output

3.1 Introduction

In the previous chapter, we introduced how to evaluate public service provision, encompassing all the main intervention areas at local level. However, we could not include in the analysis an important aspect, namely the quality aspect. This chapter addresses this issue, focusing on a specific sector where the quality aspect is noticeably important, that is the water management sector. Water is a very important component of human life, but still it is not accessible to everyone and environmentally sustainable everywhere. Accordingly, on the one hand an increasing collaboration between public and private organizations is taking off to face the huge infrastructural costs and to make the service affordable for all citizens. On the other hand, the offered service needs to be in compliance with the environmental sustainability criteria not only for the people but also for the planet. In this chapter we consider wastewater treatment as one of the main services provided by the water industry, focusing in particular on the environmental impact of the plants under analysis.

Over the last decades, wastewater treatment has received growing attention worldwide as one of the relevant activities to ensure environmental sustainability. Referring to Goodland (1995, p. 3), environmental sustainability "seeks to improve human welfare by protecting the sources of raw materials [food, water, air and energy] used for human needs and ensuring that the sinks for human wastes are not exceeded, in order to prevent harm to humans". More precisely, environmental sustainability is defined as "a set of constraints on the four major activities regulating the scales of the human economic subsystem: the use of renewable and non-renewable resources on the source side, and pollution and waste assimilation on the sink side" (ibidem, p. 10). The same concept has been proposed by the Organisation for Economic Co-operation and Development (OECD) in the OECD Environmental Strategy for the First Decade of the 21st Century (OECD, 2001, p. 6). The document considered "assimilation", one of the four specific criteria for environmental sustainability, as "the releases of hazardous or polluting substances into the environment shall not exceed their assimilative capacity". Obviously, the stated criterion is intrinsically linked to wastewater treatment and the water and nitrogen cycles are directly involved in the definition of environmental sustainability: since the main goal of wastewater treatment is to remove nitrogen and other pollutants from the ingoing water, this activity is extremely relevant on the sink side.
It is worth recalling that Environmental Sustainability is one of the three pillars of Sustainable Development and its key role is universally acknowledged[1]. Even the 2030 Agenda for Sustainable Development (UN General Assembly, 2015) confirms the triple bottom line (social, environmental and economic) approach and it defines 17 Sustainable Development Goals (SDGs) to be implemented and achieved by 2030. In this context, the improvement of water quality is considered a necessary step to ensure availability and sustainable management of water and sanitation for all (Goal 6, ibidem): water quality has to be improved "by reducing pollution, eliminating dumping and minimizing release of hazardous chemicals and materials, halving the proportion of untreated wastewa-

[1]During the World Summit on Sustainable Development in 2002 (UN General Assembly, 2002) the three pillars of Sustainable Development were identified by the words People (social sustainability), Planet (environmental sustainability) and Prosperity (economic sustainability) (see also Moldan et al., 2012).

ter and substantially increasing recycling and safe reuse globally" (Target 6.3., ibidem). Among the pollutants to be removed from water after the treatments, nitrogen is considered the most relevant one. Eutrophication, reduction of crop quality, pollution of groundwater and death of aquatic life are part of the fallout of an excessive presence of nitrogen.

The main objective of this chapter is to assess the environmental efficiency of wastewater treatment plants (WWTPs) by including in the analysis the residual nitrogen in the outgoing water. Looking at the wastewater treatment plant as a production process, the presence of nitrogen in the treated water can be considered an undesirable output. In the water sector efficiency literature, few papers deal with undesirable outputs (De Witte and Marques, 2010; Hernández-Sancho et al., 2012; Molinos-Senante et al., 2014a, 2015b; Picazo-Tadeo et al., 2008) and none of them are related to the quality of the outgoing water in terms of leftover pollutants. In line with part of the existing literature on undesirable output, the WWTP performance is evaluated using a Non-radial Directional Distance Function (NDDF) approach. As the WWTPs under analysis exhibit variable returns to scale, the Kuosmanen technology is considered (Kuosmanen, 2005) and a vector directional distance function is proposed. The efficiency scores are computed by solving a DEA-like program whose objective function is the weighted sum of the vector function components and the corresponding weights are determined by the Analytic Hierarchy Process (AHP). The integration between NDDF and AHP allows to generalize the framework proposed by Zhou et al. (2012) and by Adler and Volta (2016).

The described methodology is used to evaluate 96 WWTPs located in Tuscany (Italy). According to the different parameter specifications of the NDDF model, different efficiency scores are computed. The obtained results are then used to identify the variables affecting the efficiency. On the basis of the WWTP features, it is evident that the facilities' capacity, the percentage of wastewater discharged by industrial and agricultural activities and the level of compliance with the pollutant concentration threshold set by the Italian legislator exhibit an incontrovertible impact on the WWTP performance. The policy implications of the findings are mainly twofold: firstly, the WWTPs should exploit larger scale economies. Secondly, the water utilities, the environmental agencies and the regulators should promote inspection activities to stimulate a better functioning in particular of those plants that treat only domestic sewage and do not respect the nitrogen concentration regulatory limit.

The remainder of the chapter is organized as follows: in Section 3.2 a review of the related literature is presented and in Section 3.3 the adopted methodology is described. Then, the next three sections are devoted to the empirical analysis. More precisely, data choice can be found in Section 3.4, while the critical discussion of the performed WWTPs efficiency analysis is in Section 3.5 and in Section 3.6. In Section 3.7 the main findings of the analysis are summarized, together with some concluding remarks. Appendix B.1 includes a short description of the wastewater treatment process.

3.2 Related literature

Among the huge amount of quantitative studies on the water sector (Berg and Marques, 2011; Worthington, 2014), the wastewater treatment plant efficiency analysis has gained growing attention in recent years, i.e. starting from the 2000s (for a review, Fuentes et al., 2015). In this strand of literature, Data Envelopment Analysis (DEA) is the most used technique: it can manage a multiplicity of inputs and outputs and it does not require the selection of a specific functional form, thus resulting useful to estimate different model specifications, e.g. non-radial DEA (Hernández-Sancho et al., 2011a; Molinos-Senante et al., 2014b), DEA with uncertainty (Sala-Garrido et al., 2012a), DEA metafrontier approach (Sala-Garrido et al., 2011), Malmquist Productivity Index (Hernández-Sancho et al., 2011b; Molinos-Senante et al., 2015a). Moreover, several studies propose a further assessment of the WWTP environmental impact by means of a second stage analysis, to detect the effects of specific WWTP features on the efficiency. The most common practice is to perform non-parametric tests such as the Mann-Whitney U test and the Kruskal-Wallis test, since they do not need the normal distribution assumption of the efficiency scores (e.g. Hernández-Sancho and Sala-Garrido, 2009; Molinos-Senante et al., 2014a; Sala-Garrido et al., 2012a). In line with this strand of literature, the present chapter provides new evidence on the WWTP efficiency assessment, providing also a second stage analysis. Despite the great environmental impact of the wastewater treatment process, few papers take into account the sustainability aspects. This lack of consideration could lead to biased estimates in the performance assessment: those utilities that devote more resources to increase their environmental sustainability are penalized and turn out to be less efficient compared to those that *ceteris paribus* spend less. To address this issue in the production efficiency analysis, most studies introduce an undesirable output: it refers to those outputs whose increase may not be desirable. In the water sector performance assessment, few DEA papers deal with the undesirable output. With respect to other fields of application such as energy and cement sector, the notion of undesirable output has been conceived with a broader meaning. Looking at the various contributions, the undesirable outputs encompass unintended bad consequences (or negative externalities) which can be largely attributed to the production process, given the fact that producing good outputs is accompanied by the production of bads (Färe et al., 2014). More precisely Picazo-Tadeo et al. (2008) consider as non-desirable output the unaccounted-for water losses, De Witte and Marques (2010) and Hernández-Sancho et al. (2012) use the water losses, Molinos-Senante et al. (2016, 2015b) and Romano et al. (2017) introduce variables representing the lack of service quality such as the value of penalties, the number of complaints, the number of unplanned interruptions and the number of connected water service properties with water pressure below a reference level. Concerning the WWTPs, only Molinos-Senante et al. (2014a) deal with undesirable output by considering the CO_2 emission resulting from the WWTP activity. Except for this study, the environmental impact issue has not been addressed: as far as the authors know, none of the studies on the efficiency analysis considers outgoing water pollution. This chapter contributes to the literature addressing this gap and considering the nitrogen left in the water after the treatment as undesirable output (for more details Section 3.4 and the Appendix).

From a methodological point of view, there are several ways to model undesirable outputs, depending on how the production technology process has been formalized. Following Dakpo et al. (2016), a first approach considers the undesirable outputs as inputs and implicitly assumes its strong disposability. To mitigate this unrealistic assumption, a second strand of approaches considers undesirable outputs under the null-jointness and weak disposability assumption (see for all Färe et al. (1989) and for more recent contributions Färe et al. (2014), Adler and Volta (2016) and Färe et al. (2016)). However, in the attempt to better understand the role of undesirable outputs for some production processes, alternative approaches have been recently developed. They basically rely on the presence of some inputs directly responsible for pollution and to the possibility of identifying two distinct technologies, one related to the production of desirable outputs and the other specifically taking into account how certain inputs generate pollution. In this line, Murty (2010) defines undesirable output as a by-product *incidental output*; Hampf and Rødseth

(2015) assume the weak-G disposability on inputs and outputs; Sueyoshi and Goto (2012a,b) identify two different notions of disposability, the natural and the managerial one, to describe the managers' response to the environmental regulations by exploiting the presence of two sub-technologies[2]. The presence of two different technologies appears particularly suitable for the energy, petroleum and cement sector. By contrast, looking at the WWTPs' activity (see also Section 3.3), there is no input which can be directly associated with the chosen undesirable output, the nitrogen left in the water, and it is not possible to separate the production process into two distinct technologies. Therefore, in compliance with the above mentioned second strand of the literature, the undesirable output is modelled by assuming null-jointness and weak disposability. Regarding the efficiency assessment with undesirable output, different types of models have been used (see in particular the recent surveys by Dakpo et al. (2016) and Liu et al. (2016)). Among them, it is worth mentioning the standard DEA model, the slack-based DEA model and its extension, the model based on Russell index, the network DEA model and the Directional Distance Function (DDF) model. This latter one occupies a prominent role since it allows simultaneously for desirable output expansion and input/undesirable output contraction (see for example Picazo-Tadeo et al. (2005), Zhang and Choi (2014) and references therein): it is referred to as radial DDF if there is a proportional adjustment of the variables (Chambers et al., 1996, 1998; Chung et al., 1997; Färe and Grosskopf, 2004). Referring to the undesirable output specifically in the water sector, the DDF is considered a very suitable approach and hence it is the most developed method (e.g. Molinos-Senante et al., 2014a, 2016, 2015b; Picazo-Tadeo et al., 2008, 2011). As Fukuyama and Weber (2009) underline, the radial DDF may overestimate the efficiency when there exist non-zero slacks. To overcome this problem, several authors propose a non-radial DDF approach where slacks are directly incorporated in the efficiency measures (e.g. Barros et al., 2012; Cheng and Zervopoulos, 2014; Färe and Grosskopf, 2010; Fukuyama and Weber, 2009). A formal definition of the non-radial Directional Distance Function method (NDDF), together with several environmental indexes, is given in Zhou et al. (2012). With a similar approach, Adler and Volta (2016) suggest an economic environmental directional distance function with variable returns to scale. Going further with the NDDF approach,

[2]For a detailed discussion on the role of the undesirable output in the production process and how it can be included in a non-parametric efficiency analysis see for all Førsund (2008), Dakpo et al. (2016) and references therein.

in the present chapter the non-radial directional distance function is conceived as a vector function whose components are the scaling factors associated with the reduction of inputs, the good output expansion and the reduction of the undesirable output. With a standard technique of vector optimization, a solution is found by choosing a proper scalarization (see for example Pomerol and Barba-Romero, 2000). More precisely, the vector objective function is replaced by the normalized weighted sum of its components and the weights are analytically defined, according to the Analytic Hierarchy Process (AHP) (Saaty, 1977, 1990). A NDDF approach is then integrated by the AHP; the model proposed by Zhou et al. (2012) can be seen as a particular case. There, the normalized weight vector is chosen by "assigning the same importance" to the set of inputs, the set of outputs and the undesirable output; the same happens among the input and good output variables (see also Section 3.3). It is worth underlining that with respect to the current analysis, although several integrated DEA/AHP models have been proposed in the recent literature (see for example Pakkar (2015) and references therein), they address different issues and they serve different purposes.

3.3 Methodology

3.3.1 The WWTP production technology

The overall production process of a WWTP is characterized by a very high environmental impact and therefore the efficiency assessment of a plant cannot be separated from its sustainability performance evaluation. Therefore, the environmental production technology has to take into account inputs, good (desirable) outputs and bad (undesirable) outputs. Inputs are described by vector $x = (x_1, ..., x_N) \in \mathbb{R}_+^N$, while good and undesirable outputs are represented by $y = (y_1, ..., y_M) \in \mathbb{R}_+^M$ and $b = (b_1, ..., b_J) \in \mathbb{R}_+^J$ respectively. The environmental production technology is characterized by the following set $T = \{(x, y, b) : x \text{ can produce } (y, b)\}$ or alternatively $P(x) = \{(y, b) : (x, y, b) \in T\}$. Regarding inputs and good outputs, the environmental production technology satisfies the standard axioms of production theory (for further details see Färe and Grosskopf, 2003): i) inactivity is always possible, i.e., $(0, 0, 0) \in T$; ii) finite amount of inputs can produce only finite amount of outputs; iii) T is convex; iv) good outputs are strongly disposable, i.e., if a given amount of inputs can produce a certain level of outputs, even a smaller quantity of outputs can be produced.

According to a very standard approach (see for all Färe et al., 1989, 2014), in the present chapter, the undesirable output is considered as an unintended "by-product" of the production process and the technology is assumed to verify the following assumptions: (i) null-jointness, i.e., if $(x, y, b) \in T$ and $b = 0$, then $y = 0$; (ii) weak disposability of undesirable outputs, i.e., if $(x, y, b) \in T$, then $(x, \theta y, \theta b) \in T$, with $\theta \in [0, 1]$. Roughly speaking, null-jointness implies that there is no possibility to eliminate the undesirable outputs without stopping the good output production. Weak disposability states that a proportional reduction of undesirable outputs is possible only if it is accompanied by a corresponding proportional reduction of good outputs (for further discussion, see Section 3.4).

Looking at the wastewater treatment production, there is no empirical evidence allowing the description of the process by means of a specific functional form; therefore, non-parametric approaches are the most developed in this framework and, among them, DEA models occupy a prominent position. The technology set is then described by means of inequality and equality constraints which characterize DEA models. Moreover, preliminary analysis on the present data set show that the production processes of wastewater treatment plants exhibit variable returns to scale[3]. Following Kuosmanen (2005), the production technology can be then described as follows

$$
\begin{aligned}
T_l = \{(x, y, b) : & \sum_{k=1}^{K} \lambda^k y_m^k \geq y_m, \ \ \forall m \quad \sum_{k=1}^{K} \lambda^k b_j^k = b_j, \ \ \forall j \\
& \sum_{k=1}^{K} (\lambda^k + \mu^k) x_n^k \leq x_n, \ \ \forall n \ \ \sum_{k=1}^{K} (\lambda^k + \mu^k) = 1 \quad \lambda^k, \mu^k \geq 0, \ \ \forall k\}
\end{aligned}
\tag{3.1}
$$

where k is the number of observed production units, i.e. DMUs (in the specific context of the present chapter, WWTPs); y_m^k is the m-th desirable output produced by the k-th DMU and y_m is the m-th desirable output produced by the evaluated DMU. Similarly, b_j^k (x_n^k) is the j-th undesirable output (n-th input) produced by the k-th DMU and b_j (x_n) is the j-th undesirable output (n-th input) associated with the evaluated DMU; $\lambda^k + \mu^k$ represents the intensity weights for constructing the convex combinations of the observed DMUs. Strong disposability of inputs and good outputs are formalized by the inequality constraints and weak disposability of bad outputs is described by the equality constraints re-

[3]Regarding the scientific debate about the proper way of modelling weak disposability and variable returns to scale see also Chen and Ang (2016).

lated to b.

3.3.2 The Analytic Hierarchy Process/Non-radial Directional Distance Function approach

Taking into account the Kuosmanen technology, the environmental efficiency analysis of WWTP is performed by introducing the following vector non-radial distance function:

$$\vec{VD}(x, y, b) = \sup\{(\beta_x, \beta_y, \beta_b) : (x, y, b) + g\ \mathrm{diag}(\beta_x, \beta_y, \beta_b) \in T_l\} \quad (3.2)$$

where $\beta = (\beta_x, \beta_y, \beta_b) = ((\beta_{x_n})_{n=1}^N, (\beta_{y_m})_{m=1}^M, (\beta_{b_j})_{j=1}^J)$ is the scaling vector function. More precisely, β_{y_m} is the scaling factor of output m, β_{b_j} and β_{x_n} represent the scaling factor of the j-th undesirable output and n-th input respectively; $g = (g_x, g_y, g_b) = ((g_{x_n})_{n=1}^N, (g_{y_m})_{m=1}^M, (g_{b_j})_{j=1}^J)$ is the explicit directional vector in which the input-output combination will be scaled (Zhou et al., 2012). For each DMU, the value of the vector non-radial distance function can be obtained by solving the following DEA-like vector maximization problem

$$\begin{aligned}
\max \quad & \beta = (\beta_x, \beta_y, \beta_b) & & \\
s.t. \quad & \sum_{k=1}^{K} (\lambda^k + \mu^k) x_n^k \leq x_n + g_{x_n}\beta_{x_n}, & & \forall n \\
& \sum_{k=1}^{K} \lambda^k y_m^k \geq y_m + g_{y_m}\beta_{y_m}, & & \forall m \\
& \sum_{k=1}^{K} \lambda^k b_j^k = b_j + g_{b_j}\beta_{b_j}, & & \forall j \\
& \sum_{k=1}^{K} (\lambda^k + \mu^k) = 1 & & \\
& \lambda^k, \mu^k \geq 0, & & \forall k \\
& \beta_{x_n} \geq 0, \beta_{y_m} \geq 0, \beta_{b_j} \geq 0 & & \forall n, \forall m, \forall j
\end{aligned} \quad (3.3)$$

The above general formalization can be differently specified according to the form of β and g and hence Problem 3.3 can be seen as a general framework encompassing some relevant directional distance function models. In the present analysis, the following specifications are considered: i) $\beta = (\beta_x)$ and $g = (-x, 0, 0)$, ii) $\beta = (\beta_x, \beta_y)$ and $g = (-x, y, 0)$, iii)

$\beta = (\beta_x, \beta_y, \beta_b)$ and $g = (-x, y, -b)$. In the first case, the model allows for the reduction of inputs, while in the second case it deals with simultaneous input reduction and output expansion. The third specification of β and g takes into account the input and bad output reduction together with the output expansion. The vector formulation emphasizes that improvements for inefficient DMUs can be suggested through different directions; the non-radial approach is therefore taken to its extreme. However, from a mathematical point of view, a vector maximization problem has multiple non-dominated solutions and, among them, it is necessary to identify the most suitable ones with respect to the environmental efficiency analysis of the DMUs [4]. Following a very standard approach (see Pomerol and Barba-Romero, 2000), the vector maximization problem can be solved by maximizing a scalar function of the type $w^T\beta$ where $w = ((w_{x_n})_{n=1}^N, (w_{y_m})_{m=1}^M, (w_{b_j})_{j=1}^J)$ is the normalized weight vector, that is $\sum_{n=1}^{N} w_{x_n} + \sum_{m=1}^{M} w_{y_m} + \sum_{j=1}^{J} w_{b_j} = 1$ and $w_{x_n} \geq 0, w_{y_m} \geq 0, w_{b_j} \geq 0$, $\forall n, m, j$. Once the set of weights is chosen, the environmental efficiency score of each DMU is obtained by solving the following scalar problem

$$
\begin{aligned}
\max \quad & w^T\beta & \\
s.t. \quad & \sum_{k=1}^{K}(\lambda^k + \mu^k)x_n^k \leq x_n + g_{x_n}\beta_{x_n}, & \forall n \\
& \sum_{k=1}^{K}\lambda^k y_m^k \geq y_m + g_{y_m}\beta_{y_m}, & \forall m \\
& \sum_{k=1}^{K}\lambda^k b_j^k = b_j + g_{b_j}\beta_{b_j}, & \forall j \\
& \sum_{k=1}^{K}(\lambda^k + \mu^k) = 1 & \\
& \lambda^k, \mu^k \geq 0, & \forall k \\
& \beta_{x_n} \geq 0, \beta_{y_m} \geq 0, \beta_{b_j} \geq 0 & \forall n, \forall m, \forall j
\end{aligned}
\tag{3.4}
$$

[4]DEA vector optimization problems have been introduced even in other different contexts. Among them, an interesting contribution is the one by An et al. (2016) where a Nash bargaining approach is proposed.

In this regard, the non-radial directional function model proposed by Zhou et al. (2012) (see Problem 12 p. 629) can be seen as a particular case of Problem 3.4 [5].

As regards the weights' assignment, in the authors' opinion, the decision maker (in the present analysis the water utility) should be allowed to determine a set of weights according to the specific features of the evaluated production process. To address this issue, the Analytic Hierarchy Process appears a valuable tool. AHP has been developed by Saaty at the end of the seventies (Saaty, 1977) and later widely applied in different fields and contexts. In the present case, the AHP model is able to define a set of weights w which take into account the preferences of the decision maker (DM) on the relative importance of inputs, good and bad outputs. The DM is asked to establish if inputs are more important than good outputs and to define the intensity of such importance. The comparison yields a number which is determined on the basis of the standard AHP scale, from 1 (Equal Importance) to 9 (Extreme Importance).

Similar judgements have to be given between inputs and bad outputs and between good and bad outputs. Therefore, the following 3 by 3 matrix A is obtained

$$\begin{array}{c} \\ x \\ y \\ b \end{array} \begin{array}{c} \begin{array}{ccc} \quad x & \quad y & \quad b \end{array} \\ \begin{pmatrix} a_{xx} & a_{xy} & a_{xb} \\ a_{yx} & a_{yy} & a_{yb} \\ a_{bx} & a_{by} & a_{bb} \end{pmatrix} \end{array}$$

Clearly $a_{ij} > 0$, $a_{ii} = 1$ and $a_{ij} = \frac{1}{a_{ji}}\ \forall i, j$. A is called pairwise comparison matrix; the set of global weights $w = (w_x, w_y, w_b)$ is the normalized eigenvector associated with the dominant eigenvalue of matrix A (see for all Saaty, 1990). Going on with the pairwise comparisons among inputs, the relative weights of input variables are determined. More precisely, according to the DM preferences a pairwise comparison matrix I is constructed. The normalized eigenvector r associated with the dominant eigenvalue of matrix I represents the relative weights of input variables and hence the weight associated with the n-th input is defined as $w_{x_n} = w_x r_n$. The weights for desirable and undesirable output are computed following the same procedure. Furthermore, to detect the inconsistency of the DM preferences, the following Inconsistency Ratio is com-

[5]Zhou et al. (2012) deeply investigate the relationship among their models and the most relevant distance function models (radial and non-radial) in the recent literature. Their considerations still apply for the present chapter.

puted: $IR = \dfrac{\alpha_{\max} - n}{(n-1)CRI}$, where $\alpha_{\max}$ is the dominant eigenvalue of the pairwise comparison matrix, n is the dimension of the matrix and CRI represents the coefficient of random inconsistency which is computed by calculating $\dfrac{\alpha_{\max} - n}{(n-1)}$ for randomly filled reciprocal matrices (see Saaty, 1990). If $IR > 0.1$, then inconsistency occurs and the DM has to revise her/his judgement.
Following the AHP approach, the system of weights suggested by Zhou et al. (2012) can be obtained by considering pairwise comparison matrices whose entries are all ones.

Once the set of weights are determined, Problem 3.4 is solved for each DMU; the higher the optimal value, the lower the efficiency level of the evaluated unit. In line with Zhou et al. (2012), the obtained scores are then used to construct normalized efficiency indexes, where 1 corresponds to the best performance and 0 to the worst. Obviously, according to the specification of g and β different models can be considered and then different water efficiency performance indexes (WPI) can be constructed. The present analysis deals with the indexes presented in Table B.1.

Table 3.1: Indexes

	Input model	Input/Good Out. model	Input/Good/Bad Out. model
β	$\beta = (\beta_x)$	$\beta = (\beta_x, \beta_y)$	$\beta = (\beta_x, \beta_y, \beta_b)$
g	$g = (-x, 0, 0)$	$g = (-x, y, 0)$	$g = (-x, y, -b)$
Index	$WPI_1 = 1 - \sum_{n=1}^{N} w_{x_n}\beta_{x_n}$	$WPI_2 = \dfrac{1 - \sum_{n=1}^{N} w_{x_n}\beta_{x_n}}{1 + \sum_{m=1}^{M} w_{y_m}\beta_{y_m}}$	$WPI_3 = \dfrac{1 - \left(\sum_{n=1}^{N} w_{x_n}\beta_{x_n} + \sum_{j=1}^{J} w_{b_j}\beta_{b_j}\right)}{1 + \sum_{m=1}^{M} w_{y_m}\beta_{y_m}}$

3.3.3 Identifying WWTP efficiency explanatory variables

In line with a growing part of the literature, a second-stage analysis is performed to identify whether there is a relationship between some WWTP features and the efficiency scores obtained as described in the previous section. Regression analysis is one of the most common methodological approach, but its application presents few drawbacks: among them, there are the misspecification of the model because of omitted variables that should have been introduced rather in the first stage (Hernández-

Sancho et al., 2011b) and inaccurate results that might arise from serial correlation between the error term and the covariates in the second stage (Simar and Wilson, 2007). Therefore, in this chapter a different approach is preferred and it is applied into two steps: first of all, the WWTPs have to be categorized into groups by different operational factors that could affect the WWTP performance; then, a test is performed to assess whether or not there is statistically significant difference between/among groups according to the explanatory factor under scrutiny. As the WWTP sample does not satisfy all the necessary assumptions to apply parametric and statistical tests such as the *t-test* or the *analysis of variance*-ANOVA (Hernández-Sancho et al., 2011b; Molinos-Senante et al., 2014b), the corresponding non-parametric test is performed: the Mann-Whitney U test applies for two groups, while the Kruskal-Wallis test for three groups or more (see e.g. Kruskal and Wallis (1952) and Ruxton and Beauchamp (2008) for further details). The null hypothesis states that the groups/samples originate from the same population, while the alternative hypothesis asserts that they originate from other populations. The null hypothesis is rejected for a p lower than or equal to 0.05: if this is the testing result, it is possible to conclude that the factor under investigation does affect WWTP efficiency.

3.4 Data

The empirical analysis involves 96 wastewater treatment plants located in Tuscany and controlled by Acque SpA, a public-private utility entrusted in 2002 with water services in the so called "Basso Valdarno" river basin in the Pisa province. Data are provided by Acque Spa and refer to 2014. The data grid for this study has been constructed with the support of the Tuscan water authority staff and the technical staff of Acque SpA and Ingegnerie Toscane. The data have been gathered by a team of engineers and their consistency has been double-checked by Acque management and researchers. In compliance with the basic DEA requisites, the sample consists of a group of homogeneous WWTPs to be compared: the units under analysis refer to those plants that have costs both for the water treatment and for the sludge process and they have been refined by means of a preliminary outlier detection analysis.

As a fundamental step in the efficiency assessment, the variables have been selected not only according to the related literature, but also according to the opinion of the engineers and the data availability (e.g. Fuentes et al., 2015). Before defining the variables, it is worth pointing out that as concerns the input and output choice, a selection screening process has been executed as proposed in Golany et al. (1994): in particular, the correlation analysis between pairs of factors turns out to be useful to identify redundant variables and then to increase the discriminatory power of the DEA method. Table 3.2 presents the descriptive statistics of the variables introduced as follows.

Input. According to the mainstream literature, costs for the wastewater treatment functioning are considered as inputs. They can be taken both at an aggregate level as total costs (e.g. Da Cruz et al., 2012; Molinos-Senante et al., 2014a) or at a disaggregate level (e.g. De Witte and Marques, 2012; Hernández-Sancho et al., 2011a,b; Hernández-Sancho and Sala-Garrido, 2009; Molinos-Senante et al., 2014b; Sala-Garrido et al., 2012b). In this analysis three different cost items have been identified: (x_1) materials and energy costs; (x_2) staff and maintenance costs; (x_3) sludge transport and disposal costs: they all are expressed as €/year.

Desirable output. Looking at the literature on WWTP efficiency evaluation, basically two approaches can be identified for the output choice. In the first one, the volume of treated/delivered water and/or the population served are considered as output. The papers by e.g. De Witte and Marques (2010), Picazo-Tadeo et al. (2011), Da Cruz et al. (2012), De Witte and Marques (2012) belong to this first strand of literature. Alternatively, outputs can be chosen among the eliminated contaminants and the quantity of pollutants removed to value the production of a plant or as the difference between the pollution level in the influent and effluent, namely Net Environmental Benefits (see for example Fuentes et al., 2015; Hernández-Sancho et al., 2011a,b; Hsiao et al., 2007; Molinos-Senante et al., 2015a; Sala-Garrido et al., 2012a, 2011, 2012b). After the preliminary screening process and in compliance with the first strand of the literature, the treated water, expressed in m^3, is chosen as output (y_1). Considering in addition the main WWTP competences and the engineering expertise, a second output (y_2) has been selected, the Kg of removed sludge; actually it is by far the largest removed constituent (for further details see the Appendix B.1).

Undesirable output. As already pointed out in Section 3.2, there are no efficiency analysis papers dealing with WWTP water pollution as an undesirable output. Once the wastewater enters in the treatment plant, it is characterized by the presence of several constituents. Even though one of the main objectives of a WWTP should be the removal of as many contaminants as possible and to get the water purified for further reuse, it is almost impossible to remove them completely and so they are still present in the ongoing wastewater. The higher the pollutants in the outgoing wastewater, the higher the negative impact on the environment. Among the constituents, nitrogen is one of the most preeminent pollutant and its relevance is widely acknowledged in the related literature (see e.g. Lorenzo-Toja et al., 2015): accordingly, this chapter introduces as undesirable output the quantity of nitrogen which remains in the outgoing wastewater. It is worth pointing out that the chosen undesirable output can be seen as a bad externality and as an unintended by-product of the production process in the sense of Färe et al. (2014), but it is far from the definition of by-product given by Murty (2010). Actually, looking at the data and at the treatment process, the assumptions of null-jointness and weak disposability are fulfilled, ruling out other approaches proposed in the literature to model the undesirable outputs[6].

Table 3.2: Descriptive statistics

	Inputs			Desirable outputs		Undesirable output
	Materials+ Energy	Staff+ Maintenance	Sludge transport+ disposal	Treated water	Removed sludge	Residual Nitrogen
	€	€	€	m^3	Kg	Kg
Mean	41,894.25	13,556.09	33,190.35	402,233.20	503,315.50	5,557.39
Std. Dev.	86,448.89	23,710.57	67,347.96	942,951.40	757,611.10	11,745.93
Min	1,270.81	375.19	316.90	1,515.00	2,000.00	70.75
Max	529,684.70	145,919.00	447,154.00	6,234,272.00	4,659,130.00	78,551.83

3.5 The WWTPs performance assessment

3.5.1 Model set-up

As explained in Section 3.3.2, the choice of the normalized weight vector w is a key element of the analysis: in the following, two sets of weights are used. The first one is constructed by assigning the same importance to the three groups of variables (inputs, good and bad outputs) and the

[6]In the presented WWTP framework, there are no inputs which are specifically related to the "production" of undesirable output and therefore one of the five attributes for the by-product technology "à la Murty" fails to be verified. Therefore, the technology proposed in Murty et al. (2012) cannot be used in the present context. The authors are grateful to an anonymous referee for giving the opportunity to better clarify this important and debated aspect (for a broader overview, see Dakpo et al., 2016).

same applies inside each group (see also Section 3.3.2). The obtained weight vector coincides with the one proposed by Zhou et al. (2012). Referring to the second set, different pairwise comparison matrices are taken, following a discussion with the water utility staff. More precisely, the importance of the undesirable output is judged "very strong" with respect to good output (input). On the input side, the first two inputs share the same level of importance and they are strongly more important than the third input. Finally, with respect to the kg of removed sludge, the importance of treated water is judged very strong. Table 3.3 describes the chosen pairwise comparison matrices, the associated inconsistency ratio (IR) and the corresponding generated weights.

Table 3.3: AHP-non-radial set of weights

	Global comparison	Inputs	Good Outputs	Bad Output
Matrix	$\begin{pmatrix} 1 & 3 & 1/7 \\ 1/3 & 1 & 1/9 \\ 7 & 9 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 5 \\ 1 & 1 & 5 \\ 1/5 & 1/5 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 7 \\ 1/7 & 1 \end{pmatrix}$	
IR	0.0692	0	0	
WPI_1 Weights	$(1, 0, 0)$	$(0.455, 0.455, 0.09)$		
WPI_2 Weights	$(0.75, 0.25, 0)$	$0.75 * (0.455, 0.455, 0.09)$	$0.25 * (0.875, 0.125)$	
WPI_3 Weights	$(0.149, 0.066, 0.785)$	$0.149 * (0.455, 0.455, 0.09)$	$0.066 * (0.875, 0.125)$	0.785

Before showing the obtained results, it is worth pointing out that a further index WPI_0 has been computed in addition to the three indexes described in Section 3.3.2 and listed in Table B.1: for the sake of comparison, WPI_0 has been constructed by considering a non-radial Directional Distance Function model where the undesirable output is completely ignored. Referring to the parameter specification, the form of β and g has been set as $\beta = (\beta_x, \beta_y)$ and $g = (-x, y)$, while the equality constraint associated with the weak disposability of the undesirable output has been cancelled. The set of weights of WPI_0 coincides with the one for WPI_2.

3.5.2 Results

For both sets of weights and for the 96 WWTPs, the efficiency indexes have been computed. Table 3.4 presents in a synthetic way the main descriptive statistics of the results: the full list of the results are in Appendix B.2. The first part of the table is denoted as "Non-radial" and it refers to the set of weights proposed by Zhou et al. (2012). The second part is referred as "AHP-non-radial" and it is related to the set of weights coming from the pairwise comparison matrices of Table 3.3[7].

Table 3.4: Efficiency score for each Wastewater Performance Index

	WPI_0	WPI_1	WPI_2	WPI_3
Non-radial				
Mean	0.430	0.654	0.556	0.564
Std. Dev.	0.322	0.277	0.349	0.338
AHP-non-radial				
Mean	0.419	0.647	0.571	0.623
Std. Dev.	0.333	0.290	0.338	0.306
N^o efficient WWTPs	19	33	33	31

Not surprisingly, the number of efficient units increases as the undesirable output enters in the analysis. In fact, the wastewater treatment plants might face higher costs because of their water quality concern: the more their effort, the more efficient the wastewater treatment process and the lower the quantity of dangerous nitrogen in the outgoing water. If this aspect is not considered in the WWTP performance assessment, the "environmentally oriented" WWTPs will be penalized.
By comparing the results across the three models with the undesirable output (WPI_1, WPI_2 and WPI_3), more discriminating power is shown by the one where the contraction of both inputs and bad output and the expansion of outputs are simultaneously considered (WPI_3). Looking inside the same model specification, the number of efficient units does not change even if the normalized weight vector varies. However, the different weighting scheme affects the WWTP performance assessment in terms of efficiency scores: as it will be clarified in the next section, it influences also the explanatory variable investigation. Not surprisingly, the WPI_2 and WPI_3 efficiency scores in the AHP-non-radial case are higher than in the non-radial one: those plants who are more "environmentally" focused are valued more for keeping their water quality commitment when the undesirable output is taken into account and its importance is considered way more important than the other variables in the assessment.

[7] All models have been implemented using MATLAB 9.0 R2016a.

3.6 A step further in the WWTPs performance assessment

Once the wastewater environmental performance indexes have been obtained, a further analysis has been performed. Firstly, the variables affecting the WWTP environmental impact and their efficiency assessment have been detected. Then, the main evidence has been discussed so to give the water authorities and the decision makers additional insights for the WWTP management.

3.6.1 External factor choice

In compliance with the related literature (e.g. Hernández-Sancho et al., 2011a,b; Molinos-Senante et al., 2014a,b, 2015a) and according to the opinion of the engineers, the data availability and the WWTP main activities (for more technical details see the Appendix), the following external factors have been put under scrutiny. (1) *Age*: two groups of WWTPs have formed depending on whether the buildings are under 30 years old (that is after or before 1985). The thresholds have been set considering that the useful life of a WWTP is averagely 25/30 years, so that a plant over 30 years old is outdated. (2) Plant capacity expressed as *Population Equivalent*[8]: the facility size can be expressed in terms of per capita and per day pollution load; three groups have been defined in accordance with the Decision on Implementation Programmes (European Commission, 2007): i) less than 2000 PE, ii) between 2000 and 10,000, and iii) greater than 10,000. (3) *Sewage system*: WWTPs have been split into two groups, as they can have either a separate or a combined system. (4) *Kind of treatment*: two groups have been defined as the WWTPs in the sample use either secondary or tertiary treatment. (5) *Technologies*: referring to the wastewater treatment technology, WWTPs have been divided into two groups, as they can use or not activated sludge process. (6) *Estimated dry weather flow*: it refers to the wastewater flow occurring during the dry season when groundwater infiltration and surface runoff have a minimum influence: i) less than 100,000 m^3/year, ii) between 100,000 and 500,000 m^3/year, and iii) greater than 500,000 m^3/year. (7) *Wastewater discharged by industrial and agricultural activities* (expressed as %): the WWTP sample has been clustered in three different groups, consider-

[8] "1 P.E. (Population Equivalent)" means the organic biodegradable load having a five-day biochemical oxygen demand (BOD5) of 60 g of oxygen per day (directive 91/271/EEC).

ing separately i) those with no wastewater discharged by these activities at all, ii) those with a percentage lower than 10, and iii) those with one higher than 10. (8) *N Concentration regulatory limit*: since the detrimental effect of nitrogen in terms of eutrophication and environmental impact on plant and aquatic life is acknowledged also by the national legislator, two groups of WWTPs have been distinguished depending on whether the outgoing nitrogen concentration is below or above the limit (30 mg/L) set by the legislative decree 152/2006.

3.6.2 Results and policy remarks

Table 3.6 presents the Mann-Whitney U and Kruskal-Wallis test results and the average efficiency scores, together with the efficient WWTP percentages and the standard deviations, for each Water Performance Index and for both the weighting schemes over the WWTPs under analysis, grouped according to the described explanatory variables[9]. Before going into detail, the results suggest two preliminary considerations. First of all, when the undesirable output is considered in the efficiency assessment, there can be different test outcomes: this is the case between, on the one hand, WPI_1 and, on the other hand, WPI_2 and WPI_3. As already highlighted in Section 3.5.2, the inclusion of the undesirable output in the WWTP technology process allows those more "environmentally focused" plants to be evaluated in a fairer way and, more broadly speaking, it better depicts the overall WWTP process framework. Accordingly, the factors that might affect the performance assessment are better captured when also the undesirable output is included in the analysis. Moreover, it can be observed that there are no remarkable differences across the two different weighting scheme proposed in the current analysis. However, this might not have been the case if the water utilities would have assigned different level of importance to the inputs, good and bad outputs at the efficiency analysis stage.

Among the examined factors, the facility capacity seems to play an important role. In fact, considering as its proxy both the Population Equivalent and the dry weather flow, the efficiency means are greater for WWTPs with bigger capacity than for smaller plants and the Kruskal-Wallis test results lead to reject the equality of means hypothesis especially in the WPI_3 model. These evidences are consistent to each other and in line with other empirical applications (e.g. Hernández-Sancho

[9]Descriptive statistics and test results have been obtained using Stata 13.

et al., 2011b; Molinos-Senante et al., 2014a,b). This suggests water utilities' room for improvement in terms of unexploited economies of scale: large plants show a good hydraulic performance and their pumps achieve a high productivity. Furthermore, a higher scale of operations allows the adoption of more advanced technology and therefore a higher removal rate is obtained.

A different reasoning applies for the aspects related to the adopted technologies: the Mann-Whitney U test results do not lead to reject the null hypothesis. In fact, the distinction between secondary or tertiary, combined or separate, activated sludge process usage or not usage, does not suggest statistically significant differences and therefore these variables cannot be considered as explanatory factor for the efficiency assessment in this context. For example, the lack of significance between secondary and tertiary treatment might rely on the opposite effects exerted by high advanced treatment process (as tertiary) on costs and quality: as a matter of fact, costs grow up for the relevant capital expenditure as well as the water quality is improved. In general, these evidences are in line with other applied studies: one explanation can be related to the WWTPs sample choice. In fact, to perform an efficiency analysis, the units have to be rather homogeneous in the treatment process: looking at the specific features of the sample under analysis, this requirement is fulfilled despite the different classification. However, exploring these factors has at least two advantages: firstly, it is useful to double check the selected sample in the first stage of the analysis. Secondly, it is possible to observe the characteristics of the most efficient WWTPs: on average, there are higher efficiency scores and higher presence of efficient plants in the group that shares a tertiary, separate technology and does not use activated sludge process. Then, an interesting consideration stems from the feature related to the wastewater discharged by the industrial and agricultural activities: the efficiency score averages increase as the percentage of the efficient WWTPs increases, but the rejection of the test null hypothesis depends on the model specification. In fact, the null hypothesis is rejected for both the weighting schemes only when the undesirable output is considered: this might suggest that the efficiency assessment can be conditional upon the main target of the WWTP operators. The result is in conflict with prior literature that shows a poor performance among the plants treating sewage from factories and farms (Guerrini et al., 2016). The novelty of the result obtained in the current chapter can be attributed to the measurement of nitrogen as undesirable output in the efficiency model, mainly for two reasons. First of all, the sewage produced by some

farms and factories (mainly paper mills in the Pisa area) is poorer of nitrogen than the domestic wastewater: accordingly, the residual amount of this pollutant after the treatment process is rather low, making the plants performance better. Moreover, the plants treating sewage from factories generally turn out to be more "environmentally" oriented. Instead, referring to the year of the plant, the Mann-Whitney U test results do not enable to reject the null hypothesis in most cases. Consistently with other studies (e.g. Hernández-Sancho et al., 2011b; Molinos-Senante et al., 2014a,b), age cannot be considered as a determinant factor in the efficiency assessment.
Lastly, it is possible to observe that the two groups obtained following the concentration limit set by the Italian legislator are statistically different whenever the undesirable output is taken into account. This information provides useful insights: in fact, even if few efficient WWTPs are found among those above the set threshold, the efficiency score average is way larger in the case the plants manage to keep the outgoing concentration below the limit. Accordingly, this evidence suggests that the "environmentally focused" WWTPs benefit in terms of performance assessment rather than being damaged, despite their water quality commitment and the incurred high costs to develop a good treatment with high pollutant removal rate.

In terms of policy implications, the obtained results could provide useful suggestions for the water utilities, the environmental agencies and the regulators. For example, since "big is better" for wastewater treatment, the utilities' managers should plan to exploit larger scale economies: this would imply higher cost savings, but at the same time higher environmental standards achievement. From the point of view of the environmental agencies, the highest efficiency scores obtained by the plants serving factories and farms suggest to perform an inspection activity on small plants treating only domestic sewage. Moreover, the environmental controls concerning the nitrogen concentration regulatory limit should also be increased, so to stimulate a better functioning of those plants that do not respect the set threshold and show a lower performance level. Finally, the evidence shows the water authorities the benefits that could arise from an integrated performance assessment that penalizes WWTPs aiming only at getting cost savings and achieving poor environmental standards. The results obtained by the adoption of the "environmental performance index" might suggest water regulators a benchmarking model for the WWTPs and a yardstick competition to water utilities regulation.

Table 3.5: WPIs by explanatory factors and Mann-Whitney U/Kruskal-Wallis test results

Non-radial		WPI_1				WPI_2				WPI_3			
Explanatory factor	Total WWTPs	% Eff.	Mean	Std. Dev.	Test	% Eff.	Mean	Std. Dev.	Test	% Eff.	Mean	Std. Dev.	Test
Year													
<1985	44	36%	0.677	0.278	0.5185	36%	0.611	0.322	0.0637	34%	0.626	0.311	0.0584
≥ 1985	52	33%	0.634	0.277		33%	0.510	0.366		31%	0.511	0.353	
PE													
<2,000	57	32%	0.644	0.269	0.0536	32%	0.510	0.355	0.0577	28%	0.497	0.343	**0.0138**
2,000 - 10,000	29	31%	0.606	0.288		31%	0.560	0.324		31%	0.601	0.303	
10,000 - 150,000	10	60%	0.852	0.220		60%	0.812	0.289		60%	0.837	0.271	
Estimated Dry Weather Flow													
<100.000	63	29%	0.625	0.264	**0.045**	29%	0.497	0.342	**0.0207**	25%	0.486	0.329	**0.0027**
100,000 - 500,000	25	40%	0.651	0.307		40%	0.610	0.348		40%	0.656	0.318	
>500.000	8	63%	0.889	0.170		63%	0.861	0.227		63%	0.888	0.200	
Sewage System													
Combined	39	28%	0.608	0.280	0.2103	28%	0.487	0.356	0.0741	26%	0.491	0.340	0.0657
Separate	57	39%	0.686	0.273		39%	0.604	0.339		37%	0.613	0.330	
Level of Treatment													
Secondary treatment	92	34%	0.644	0.277	0.0885	34%	0.542	0.348	0.0885	32%	0.548	0.336	0.0704
Tertiary treatment	4	50%	0.892	0.125		50%	0.881	0.139		50%	0.918	0.106	
Technologies													
Others	6	50%	0.778	0.276	0.3619	50%	0.731	0.347	0.2726	33%	0.642	0.356	0.7348
Activated sludge	90	33%	0.646	0.277		33%	0.545	0.348		32%	0.558	0.338	
% industrial WW													
No activity	60	30%	0.628	0.267	0.1144	30%	0.509	0.341	**0.0366**	27%	0.508	0.330	**0.0102**
≤ 10%	26	40%	0.650	0.317		40%	0.599	0.366		40%	0.636	0.339	
>10%	10	60%	0.845	0.231		60%	0.826	0.262		60%	0.852	0.238	
***N* Concentration regulatory limit**													
Below (≤30 mg/L)	69	39%	0.668	0.291	0.6805	39%	0.601	0.349	**0.0419**	39%	0.631	0.333	**0.0015**
Above (>30 mg/L)	27	22%	0.619	0.239		22%	0.443	0.326		15%	0.391	0.289	

Table 3.6: WPIs by explanatory factors and Mann-Whitney U/Kruskal-Wallis test results

AHP-non-radial		WPI_1				WPI_2				WPI_3			
Explanatory factor	Total WWTPs	% Eff.	Mean	Std. Dev.	Test	% Eff.	Mean	Std. Dev.	Test	% Eff.	Mean	Std. Dev.	Test
Year													
<1985	44	36%	0.692	0.277	0.1495	36%	0.628	0.312	**0.0434**	34%	0.677	0.278	0.0953
≥ 1985	52	33%	0.609	0.297		33%	0.524	0.354		31%	0.577	0.323	
PE													
<2,000	57	32%	0.634	0.286	0.1142	32%	0.518	0.347	**0.0242**	28%	0.540	0.320	**0.0009**
2,000 - 10,000	29	31%	0.605	0.291		31%	0.586	0.306		31%	0.694	0.240	
10,000 - 150,000	10	80%	0.840	0.251		80%	0.833	0.266		60%	0.889	0.180	
Estimated Dry Weather Flow													
<100.000	63	29%	0.615	0.281	0.0777	29%	0.506	0.334	**0.0071**	25%	0.539	0.308	**0.0003**
100,000 - 500,000	25	40%	0.651	0.308		40%	0.637	0.321		40%	0.741	0.240	
>500.000	8	63%	0.883	0.200		63%	0.878	0.209		63%	0.916	0.156	
Sewage System													
Combined	39	28%	0.581	0.295	0.0659	28%	0.513	0.339	0.114	26%	0.575	0.308	0.2261
Separate	57	39%	0.691	0.280		39%	0.611	0.334		37%	0.656	0.303	
Level of Treatment													
Secondary treatment	92	34%	0.636	0.290	0.1074	34%	0.557	0.337	0.092	32%	0.609	0.305	0.0734
Tertiary treatment	4	50%	0.899	0.144		50%	0.898	0.147		50%	0.938	0.099	
Technologies													
Others	6	50%	0.788	0.304	0.3078	50%	0.751	0.350	0.2794	33%	0.674	0.358	0.8534
Activated sludge	90	33%	0.637	0.288		33%	0.560	0.336		32%	0.620	0.304	
% industrial WW													
No activity	60	30%	0.619	0.281	0.0727	30%	0.527	0.330	0.0631	27%	0.569	0.307	**0.0073**
≤ 10%	26	40%	0.637	0.327		40%	0.615	0.345		40%	0.716	0.253	
>10%	10	60%	0.857	0.220		60%	0.821	0.286		60%	0.870	0.222	
***N* Concentration regulatory limit**													
Below (≤ 30 mg/L)	69	39%	0.660	0.301	0.5191	39%	0.625	0.329	**0.005**	39%	0.719	0.262	**0.0001**
Above (>30 mg/L)	27	22%	0.612	0.261		22%	0.434	0.327		15%	0.378	0.277	

3.7 Conclusion

In this chapter, 96 wastewater treatment plants located in Tuscany (Italy) are evaluated through a novel integrated AHP/NDDF approach. The wastewater treatment plants production process is described by the following variables: material and energy costs, staff and maintenance costs and sludge transport and disposal costs are chosen as inputs; treated water and kg of removed sludge are the desirable outputs while the undesirable output is represented by the quantity of nitrogen in the outgoing wastewater. The selection of nitrogen as undesirable output is related to the environmental quality of the outgoing water and adds a new dimension to the literature on the WWTP efficiency analysis.

From a methodological point of view, this chapter goes a further step along the path traced by Zhou et al. (2012). The vector directional distance function allows a new formulation for the simultaneous reduction of inputs and bad outputs together with the expansion of good outputs. According to the specification of the non-radial distance function and the explicit directional vector, different combinations of inputs/outputs can be analysed and thus different efficiency indicators can be constructed. The normalized weight vector is selected by taking into account the decision makers' preferences (water utility managers, water authorities) and following the AHP methodology. In this regard, the suggested model encompasses the one proposed by Zhou et al. (2012). In the empirical analysis, two different sets of weights are specified presenting thus two models. In the first case (Non-radial model), the associated weight vector coincides with the one in Zhou et al. (2012). In the second case (AHP-non-radial model), the set of weights is constructed starting from the water utility staff suggestions. The computed environmental indexes differ across the two models, although the efficient units are the same.

The environmental efficiency is explained by means of several variables related to the technical features of the WWTP. Irrespective of the model specifications, the population equivalent size and the estimated dry weather flow have a significant impact on the WWTP performance. This evidence represents a clear indication for water utilities in term of WWTPs' size. On the other hand, whenever the expansion of outputs and the contraction of undesirable output are allowed, the efficiency scores for both the Non-radial and the AHP-non-radial model are affected by the percentage of the discharged industrial wastewater and by the plants' ability to respect the legal nitrogen concentration threshold. From the environmental agency side, the introduced performance indexes suggest inspection activities to control those plants that treat only domestic sewage and/or do not respect the nitrogen concentration regulatory limit.

In the present application, WWTP efficiency is addressed together with the environmental sustainability issue, specifically referring to the quantity of nitrogen in the outgoing water. The analysis of the environmental quality of the treated water might take into account other relevant residuals such as phosphorus, pharmaceutical pollutants, toxic metals and therefore further undesirable outputs might be chosen. In this context, the proposed methodology might be very promising for further inspections on the environmental efficiency of the wastewater treatment plants.

This chapter deals with the service provision analysis of a particular sector, namely the water management. Specifically, it addresses the analysis of the wastewater treatment services: plants are valued more if more "environmentally" focused and committed to water quality, providing more sustainable and better quality services. However, the measurement of service provision is only one side of the coin. The other is the evaluation of the way resources are spent: Chapter 4 and 5 explore this complementary and crucial topic in public sector management.

Part II
Public spending

Chapter 4

Global public spending efficiency in Tuscan municipalities

4.1 Introduction

In the first part of this dissertation, service provision has been under analysis. However, the necessity of combining public service delivery with the containment of public spending calls for further investigation, checking whether public expenditure is managed in compliance with the principles of efficiency and effectiveness. This chapter offers an analysis on municipal spending efficiency, while the next one investigates the effectiveness of public resources. Specifically, while Chapter 2 aimed to measure public service provision embedding in a synthetic index several municipal service indicators, this chapter deals with the efficiency assessment of municipal expenditure, linking resource management with the provision side . Accordingly, a different methodology is proposed to addressed the different policy-relevant research question.

After decades of research, local government efficiency evaluation is still at the centre of political and academic debate, in the public sector literature, and even more, in the public administration and management literature. Economic performance measurement and comparison at each government level remain a relevant issue in the current agenda, being a recurrent theme during the evolution of public sector management along its three different phases (Osborne, 2006). The first one goes from the late 19th century through the late 1970s/early 1980s; at least in the majority of European countries, the state was supposed to satisfy all the social and the economic needs of its citizens. Recalling a very famous sentence, this should have been done “from the cradle to the grave”. “Administration”, “bureaucracy” and “public service provision” characterized the activity of the public administration in that period. The second phase can be associated with the “New Public Management (NPM)” paradigm (Hood, 1991). In this era, “market”, “managerialism”, “input and output control”, “performance evaluation” got a foothold in public administration (Hughes, 2003); both theoretical and political debates faced the necessity of combining public service provisions with the containment of public spending. So, since the beginning of the 1990s, “efficiency”, “effectiveness” and “quality service” have become the keywords of public sector management (Keating, 1998). From the late 1990s, the New Public Management paradigm has been heavily criticized and much empirical evidence underlined its failure: this has lead to new proposals which attempt to give a more modern idea of Public Management Governance (see for example Dunleavy et al. (2006); Osborne (2006)). Even in these new contexts, performance evaluations are still considered as a key tool, something essential for policy makers’ decisions. The provision of a robust efficiency measurement and the implementation of an effective system of incentives are in the agenda of both politicians and academics (Da Cruz and Marques, 2014).

Local governments are the most involved organizations in the evaluating process; during the last years, many key public functions have been transferred from national to local authorities and hence these latter ones have increased their importance (Afonso and Fernandes, 2008; Afonso and Venâncio, 2016; Doumpos and Cohen, 2014; Lo Storto, 2016). As it is better specified in Section 4.2, there is a growing number of papers dealing with efficiency evaluations of local governments and the identification of those environmental variables which may affect efficiency. Several different aspects of local government activity have been evaluated with different techniques. The present chapter fits into this wide literature and in particular it aims at evaluating the efficiency of the Italian municipalities located in Tuscany.

To respect the budget constraints, the national government often makes cuts in transfers to regional and local governments and tries to reorganize public service supply. Referring to Italy, this subject is very relevant due to the stringent budget constraints imposed at European level, like the Stability and Growth Pact and more recently the Fiscal Compact. Of course, this strengthens the importance and the usefulness of efficiency evaluations. In Tuscany and, in general, in the Italian context, the presence of inefficiency in municipal expenditure is due to at least three aspects: the presence of much too small municipalities, the partial overlapping of functions carried out both by provinces and municipalities and the lack of a unitary management for densely populated metropolitan areas (Iommi, 2011). In this chapter the first aspect is specifically investigated: small municipalities turn out to be inefficient because they are unable to exploit scale economies in the provision of public goods and services and, as a consequence, the services they can provide are poorer and limited to essential needs. So, the issue of local governments' optimal size to settle these diseconomies is still controversial and matter of debate. In particular, Tuscany has promoted institutional and administrative reforms to overcome the presence of too many fragmented municipalities and to define appropriate territorial areas for planning and supply of public services: since the 1970s, there was awareness among scholars and regional administrators that very small municipal dimensions affected public service supply and that institutional boundaries were de facto already overcome in the everyday life of families and businesses. The 68/2011 regional law represents an example of the legislator's attempt to define the optimal municipal size to offer fundamental public services by promoting joint management and/or merger among the smallest Tuscan municipalities. In this context, expenditure efficiency analysis of Tuscan municipalities is proposed through a Data Envelopment Analysis (DEA). This chapter contributes to the literature by supplying new evidence concerning efficiency analysis of local governments and by proposing an innovative use of a composite indicator. Additionally, the obtained results can help the policy-maker to identify inefficient municipalities and to give suggestions on possible reorganizations of local governments.

The remainder of this chapter is organised as follows. Section 4.2 provides a literature review to place this research into context. Section 4.3 introduces the model specification, describing the 3-stage DEA based approach performed in the analysis. Section 4.4 presents the empirical analysis, explaining the data choice and the critical discussion of the obtained results. Finally, Section 4.5 concludes the chapter.

4.2 Literature review

Despite the fact that the measurement of efficiency in the private sector dates from the seminal contribution of Farrell (1957), the issue of local governments efficiency has been addressed just since the 1990s. The existing literature on municipal efficiency analysis can be divided into two branches (Doumpos and Cohen, 2014). On the one hand, there are numerous studies on individual public services, such as solid waste, sewage disposal, water, energy provision, hospitals, municipal savings banks, public libraries, road maintenance, fire protection, care for the elderly sector, local police services, public transportation and pre-school education (for an overview see Bönisch et al., 2011). On the other hand, there are studies that analyse global municipal efficiency for various countries: Belgium (De Borger and Kerstens, 1996; De Borger et al., 1994; Geys and Moesen, 2009), Finland (Loikkanen and Susiluoto, 2005), Norway (Borge et al., 2008), Brazil (De Sousa and Stošić, 2005), Spain (Balaguer-Coll et al., 2007, 2013; Benito et al., 2010; Cuadrado-Ballesteros et al., 2013; Prieto and ZofIo, 2001), Portugal (Afonso and Fernandes, 2006, 2008; Afonso and Venâncio, 2016; Da Cruz and Marques, 2014), Czech Republic (Št'astná and Gregor, 2015), Japan (Nakazawa, 2013; Nijkamp and Suzuki, 2009), Germany (Geys et al., 2010; Kalb et al., 2012), Greece (Athanassopoulos and Triantis, 1998; Doumpos and Cohen, 2014) and Italy (Agasisti et al., 2016; Boetti et al., 2012; Bollino et al., 2012; Lo Storto, 2013, 2016) (for earlier studies review see De Borger and Kerstens, 2000; Worthington et al., 2000). This second type of studies sometimes attempts to analyse the relationship between municipal performance and some important topics, like the relevance of municipal size, the effect of public function decentralization on the municipalities, the impact of fiscal decentralization, the influence of the effects of spatial closeness between municipalities, and other aspects. According to many authors, there is an advantage in the use of a comprehensive approach, compared to the studies focused on specific functions: it is the ability to take into account the opportunity cost perceived by the municipality in deciding the allocation of resources to different services, the possible synergies of expenditure and the quantification of the total savings of resources. Following this part of the efficiency literature, in this chapter a global public expenditure efficiency analysis of the Tuscan municipalities is performed through DEA and, as far as the authors know, this is the first application for the Tuscan region. The choice of the Tuscan framework is undoubtedly linked to its topical feature: even the Tuscan legislator has promoted institutional and administrative reforms to overcome the presence of inefficiency in municipal expenditure, in particular in relation to municipal size. For this reason, in this context specific attention is dedicated to the municipal size effect on expenditure efficiency, adding new evidences to the existing literature (see for example Bönisch et al., 2011; Doumpos and Cohen, 2014).

From a methodological point of view, there are alternative available methods for the efficiency analysis of production processes in both private and public sector. They differ mainly in the way the unknown and unobservable "efficiency frontier" is inferred from the data. These different techniques can be classified basically in two alternative approaches: the econometric and the optimization approach. The first one specifies a production function and normally recognizes that the deviation away from this given technology (as measured by the error term) is composed of two parts, one representing randomness (or statistical noise) and the other inefficiency. Among the various techniques belonging to the econometric approach, "stochastic frontier analysis" (SFA), introduced by Aigner et al. (1977), plays a central role. Following Worthington (2000), the first studies of local government cost efficiency with this approach are proposed by De Borger and Kerstens (1996), Deller et al. (1988) and Hayes and Chang (1990). Using this technique, a sizeable structure is imposed upon the data from a strict parametric form and distributional assumption, to determine the absolute economic efficiency of the units under analysis against some imposed benchmark (Dollery and Wallis, 2001). On the contrary, the mathematical programming approach seeks to evaluate the relative efficiency of one unit compared to the others. The most commonly employed version of the optimization approach is the linear programming model referred to as "data envelopment analysis" (DEA), introduced by Charnes et al. (1978), based on the concept of efficiency proposed by Farrell (1957). DEA essentially calculates the economic efficiency of a given organisation with respect to the performance of other organisations producing the same good or service, rather than against an idealised standard of performance. Given its non-parametric basis, it is possible to considerably vary the specification of inputs and outputs and not to specify a particular form. Still following Worthington, De Borger and Kerstens (1996) and Eeckaut et al. (1993) give the first contributions for the local government cost efficiency analysis with this technique. Moreover, a less-constrained alternative to DEA often employed in the analysis of public sector economic efficiency is known as "free-disposal hull" (FDH), introduced by Deprins et al. (1984) and applied to local governments for the first time by De Borger and Kerstens (1996) and by De Borger et al. (1994). The methodological literature to date provides inconclusive evidence concerning the sensitivity of local government efficiency rankings to these alternative technologies. It should be emphasised that the SFA and DEA approaches address different questions, serve different purposes and have different informational requirements (Dollery and Wallis, 2001): for these reasons, DEA and SFA should be considered as complementary methods in local public sector efficiency analysis. Recently, Da Cruz and Marques (2014) carried out a very detailed and systematic literature review of the papers published in peer-reviewed and top-ranked journals dealing with the global performance of local governments: DEA, FDH or SFA methodologies are mostly used and the data choice is strictly affected by the local governments' range. In particular, as regards the data choice, the efficiency analysis at the global level covers several areas of municipal activity: many inputs and outputs related to different municipal areas have to be considered. To globally encompass all the municipal functions in a single indicator, several ways have been proposed in the recent literature. Some authors conceive a multi input-output DEA model (see Balaguer-Coll et al., 2007; De Borger et al., 1994; Worthington, 2000), while some others aggregate the different functions by constructing a composite indicator. Regarding this latter approach, it is worth citing the contribution of Afonso and Fernandes (2008). The authors use a Total Municipal Output Indicator (TMOI) to put together different outputs (a similar approach can be found for example in Afonso et al., 2005). They assume that the TMOI depends on several economic and social variables, belonging to different policy areas. For each policy area a total municipal sub-indicator (TMSOI) must be previously computed: this indicator is calculated by centring each variable around the mean of all observations and then using an unweighted average of all variables for a policy area. Then, the TMOI is computed as the sum of all the sub-indicators. DEA analysis is then performed using as output of the model either the composite TMOI or alternatively the several sub-indicators. Another way of constructing composite indicators is based on the so-called "Benefit of the doubt" approach (see Cherchye et al., 2007). In this case, separate sub-indicators are first computed for different objects and then they are aggregated in a composite indicator by means of their weighted sum. The weights are chosen so to maximize the value of the composite indicator: in other words, they are the most favourable weights for the evaluated unit. However, in constructing the overall efficiency score, a common system of weights would be preferable; in this light, Despotis et al. (2002) suggests a procedure which remains in “the spirit” of DEA and which determines the same weights for every unit. This kind of composite indicator is often defined as DEA-like composite indicator and it is used in several different contexts, such as the assessment of the human development index and the evaluation of quality of life and well being (see for example Bernini et al., 2013; Despotis, 2005). As far as the authors know, it has not yet been applied to local government expenditure efficiency analysis: this chapter contributes to the literature by introducing the use of this DEA-like composite indicator for the computation of the global efficiency scores. Moreover, to validate the provided results and to give an interpretation of the global efficiency scores in terms of municipal expenditure composition, also a further composite indicator is proposed.

4.3 Model specification

To evaluate overall spending efficiency, a 3-stage DEA based approach is performed: first of all, individual efficiencies associated with five major municipal functions are computed; then the municipal global efficiency index is generated considering a common set of weights; finally, statistical analysis is used to assess the effect of some contextual variables on the global efficiency indicator. In the following, each stage is explained in more details.

4.3.1 Stage 1: DEA for each individual municipal function

As it has been already underlined in Section 4.2, DEA is a non-parametric technique which is particularly suitable in evaluating the efficiency of the public sector. It does not require any specific functional form of the production frontier and gives intuitive ideas to correct the found inefficiency. Through a linear programming approach, DEA constructs the efficient frontier; first of all for each unit to be evaluated, the so-called Decision Making Unit (DMU), the set of inputs and output are detected. Then DEA models analyse whether either a given output quantity is produced with minimum input (input-oriented DEA model) or the maximum output is produced with a given input quantity (output-oriented DEA model). The efficiency score varies from 0 to 1 and is determined by the ratio between the weighted sum of outputs and the weighted sum of inputs. Moreover, regarding the possibility of allowing variable returns to scale or not, two different specifications can be distinguished: the constant returns to scale DEA model (CRS) and the variable returns to scale DEA model (VRS), introduced respectively by Charnes et al. (1978) and Banker et al. (1984).

In the present analysis, Tuscan municipalities are the evaluated DMUs. As a first step, to perform a global efficiency analysis, five DEA models are run to separately assess the efficiency of the following municipal functions: "General administration" (GA), "Social Services" (SS), "Educational services" (ES), "Road maintenance and local mobility" (RM) and "Local police" (LP). Those functions have a strategic role in local government policy and occupy a prominent position in the municipal budget. The peculiarities of municipal activities suggest to use the input-oriented DEA model with variable returns to scale (VRS). [1] For the GA function, as well as the SS and the ES functions, a "one input-one output" model is used, while in the case of RM and LP functions a "one input-two output" model is chosen (for further specification, see Section 4.4.1). The obtained basic efficiency scores are then aggregated to analyse the overall spending performance, as explained in the next stage.

4.3.2 Stage 2: Aggregating for overall efficiency analysis

Regarding the overall efficiency, several preliminary considerations should be done. The global municipal spending efficiency could be evaluated by considering a DEA model with all the input and output variables detected for the non-aggregate analysis. Nevertheless, this straightforward and easy choice does not result so appealing. As DEA allows flexibility in the choice of weights on the inputs and outputs, the greater the number of included factors, the lower the level of discrimination between efficient and inefficient units: so, discrimination can be increased by being parsimonious in the number of variables. In other words, by increasing the number of inputs and/or outputs, there is automatically, by construction, an increase in efficient DMUs. This reasoning becomes very evident looking at the DEA results stemming from the municipal analysis: gradually adding a function, in the VRS model, the number of efficient municipalities increases more and more, out of a sample of 282

[1] Nevertheless, CRS model might also be estimated to provide additional information on the municipal size debate, computing the scale efficiency as the ratio between CRS and VRS efficiency scores (Cooper et al., 2011). This analysis is kept as scope for further research.

units under analysis. In fact, just considering the "General administration" function there are only 5 efficient municipalities. Considering also the function for "Educational services" the number of efficient municipalities increases at 20. Then, adding the function for "Social Services" 49 municipalities result to be efficient. Finally, the number of efficient municipalities becomes very big introducing the "Road maintenance and local mobility" function, i.e. 82 efficient municipalities, and then the "Local police" function, i.e. 107 efficient municipalities: obviously, having so many efficient municipalities is not very informative and it's quite unreasonable. In the literature, there is an open theoretical debate on this issue. From one hand, different suggested "rules of thumb" are proposed in order to achieve reasonable level of discrimination; for example, there are proposed rules in Bowlin (1998) and in Dyson et al. (2001). On the other hand, the definition of a stringent rule seems to be too rigid and useless in relation to the research needs (see, e.g., Cook et al., 2014; Cooper et al., 2011). Referring to this issue, alternative approaches have been proposed in the municipal expenditure efficiency analysis (see Section 4.2).

The solution adopted in the present analysis consists in the introduction of a new composite indicator which aggregates the efficiency scores of the single functions and more specifically by means of a weighted average of the basic efficiency scores. The weights are determined focusing on two main objectives: first of all, they have to generate a DEA-like index, so not to penalize the units under analysis; then, they have to avoid arbitrary choice, giving a common base for municipality comparison. To reach these goals, the "benefit of the doubt" approach (see Bernini et al., 2013; Cherchye et al., 2007) represents a necessary intermediate step to generate the municipal global efficiency index with a common set of weights. More precisely, for every municipality j_0 a preliminary composite indicator is constructed, solving the following maximization problem:

$$\begin{aligned} CI_{j_0} = \quad & \max_{w_{j_0}} \sum_{i=1}^{k} y_{ij_0} w_{ij_0} \\ & \sum_{i=1}^{k} y_{ij} w_{ij_0} \leq 1 \quad j = 1, ..., m \\ & w_{ij_0} \geq \epsilon \qquad\qquad i = 1, ..., k. \end{aligned} \tag{4.1}$$

where m is the number of evaluated municipalities, k is the number of the considered functions, y_{ij} represents the DEA efficiency score related to the i_{th} function of the j_{th} municipality and $w_{j_0} = (w_{ij_0})_{i=1}^{k}$ is the weight associated with the municipality j_0. The overall efficiency score is obtained by taking a weighted sum of the five non-aggregate efficiency scores and, according to Problem 4.1, it corresponds to the optimal value; the optimal weights are the most favourable for the evaluated municipality j_0. Moreover Problem 4.1 can be seen as a standard DEA model where there is a dummy input and the non-aggregate efficiency scores are seen as output variables. As it is observed in Bernini et al. (2013), the aforementioned Composite Indicator provides a different set of weights for each DMU and this prevents DMU's comparison on a common base. In this light, Despotis et al. (2002) states that a common set of weights can be determined by solving a suitable vector optimization problem and the corresponding solutions are found by solving the following minimization problem:

$$\begin{aligned}
\min_{w_i, d_j, z} \quad & t\frac{1}{m}\sum_{j=1}^{m} d_j + (1-t)z \\
& \sum_{i=1}^{k} y_{ij} w_i + d_j = CI_j \quad j = 1, ..., m \\
& d_j - z \leq 0 \qquad\qquad\quad j = 1, ..., m \\
& d_j \geq 0 \qquad\qquad\qquad\;\; j = 1, ..., m \\
& w_i \geq \epsilon \qquad\qquad\qquad\;\; i = 1, ..., k \\
& z \geq 0
\end{aligned} \tag{4.2}$$

where w_i represents the weight assigned to the i_{th} function and d_j "measures" the distance between the "collective" score and the most favourable score for the j_{th} municipality, namely CI_j which is obtained by solving Problem 4.1. By construction, d_j is non-negative for every DMU and z represents the maximum of d_j; therefore z is the distance between the "collective" score and the DEA-like composite indicator of the most penalized DMU (see also Bernini et al., 2013). As t varies from 0 to 1, different sets of common weights are determined and each of them has different meaning. In the present analysis, $t = 0$ is considered as the most suitable choice given the institutional framework outlined in the introduction of this chapter. In fact, one of the main sources of inefficient resource management has been linked to unexploited economies of scale due to the presence of many small municipalities. When $t = 0$, Problem 4.2 gives the set of common weights which minimizes z, that is, which maximizes the efficiency score of the most penalized DMU. Accordingly, even the municipalities that are expected to be the most inefficient (namely the smallest ones) cannot complain of unfairness as concerns the choice of the model. The set of common weights is then used to compute the new composite indicators attesting the municipalities' global spending efficiency.

In the local administrators' opinion, the overall efficiency of a municipality has to be evaluated through the analysis of the expenditure composition arising from its municipal balance sheet. In this light, the obtained composite indicators are compared with a second kind of global efficiency scores. Even in this latter case, the composite indicator is the weighted sum of the single functions' efficiency scores, but each function enters in the composite indicator with the same proportion that the given function has with respect to the total expenditure. This other indicator takes into account the local administrators' perception and it is also helpful in validating the first DEA-like indicator. Moreover, to construct a composite indicator with a common system of weights among all the municipalities, the single function efficiency scores are weighted according to the Tuscan mean expenditure composition. For each municipality, the comparison of the two "expenditure composition" indicators may provide suggestions to enhance efficiency.

4.3.3 Stage 3: Investigating municipal expenditure efficiency explanatory variables

Finally, an interpretation of the obtained indicators is provided by clustering the Tuscan municipalities according to the following main municipal features and consistently with the Tuscan hallmarks: size, geography, tourism degree and socio-economic structure through the local labour system classification. This analysis aims at investigating how specific municipal features affect local public expenditure management and its efficiency.

With the same purpose, a Tobit regression model is also estimated. This approach aims at evaluating the correlation between the efficiency scores and the municipal characteristics, partly complementing the cluster analysis as introduced above. The global efficiency scores are explained taking into account municipal characteristics referring to the economic, social and political context. As suggested in the recent literature (see e.g. Agasisti and Wolszczak-Derlacz, 2015; Da Cruz and Marques, 2014; Lo Storto, 2016), to properly conduct the third stage analysis, bias-corrected efficiency scores are computed using a bootstrap procedure (see Simar and Wilson, 2000, 2007) and then used as dependent variable in the regression model. Bootstrap-based inference makes bias-corrections for the efficiency score computation, but it still assumes the "separability condition". In this chapter, we show the insights stemming from the application of this technique. On the contrary, in Chapter 2 the conditional analysis has been alternatively considered, to include in one-stage the background variables avoiding the assumption of the "separability condition" (Daraio et al., 2017), so to show the application of both techniques and the consistency of their results.

4.4 Empirical application

4.4.1 Data, inputs and outputs

A fundamental step in the definition of municipal efficiency analysis regards the choice of the decision variables, both for the computation of the efficiency scores (inputs and outputs) and for the explanation of its determinants. The Italian institutional framework strongly influences the data choice, regarding both the municipal expenditure areas and their related inputs/outputs. Specifically, "General administration", "Social Services", "Educational services", "Road maintenance and local mobility" and "Local police" are considered, as they represent not only the most fundamental competencies for the municipal budget (about 73% of total current expenditure in 2011, reference year of the analysis), but also for the services provided to the citizens, detailed in the following according to the municipal balance sheet expenditure items [2].

- General administration: it provides services regarding the institutional bodies, the administrative office, the management of tax rev-

[2] Among the excluded municipal functions, a remark has to be made about the "Environmental management" function. It presents very heterogeneous expenditure items (e.g. urban services, environmental conservation services, waste disposal service), that heavily differ among municipalities according to their own characteristics. Another source of heterogeneity comes from the presence of two different taxation systems for environmental services, namely the TARSU ("*Tassa Rifiuti Solidi Urbani*") system and the TIA ("*Tariffa Igiene Ambientale*") system. Since DEA requires homogeneous units to be compared, the authors have preferred excluding this function from the current analysis.

enue, the technical office, military services registry, civil registration and electoral services, vital records and statistics.

- Educational services: it provides services regarding the nursery schools, primary and secondary education, school assistance, school transport and school meals.
- Social services: it provides services regarding childcare, kindergarten, services to minors, leisure structures, facilities and care for the most vulnerable population groups such as elderly and immigrants.
- Road maintenance and local mobility: it provides services regarding viability, traffic circulation, public lighting and public transport.
- Local police: it provides services regarding the municipal police, the commercial police and the administrative police.

In the empirical literature there is a general consensus regarding the choice of cost related observations as input (Afonso and Fernandes, 2008; Doumpos and Cohen, 2014; Kalb et al., 2012; Lo Storto, 2016): accordingly, the municipal current expenditure of each municipal area is used as input indicator, taken in non-aggregate way and expressed in absolute value. Data come from the available municipal balance sheets, published by the Home office Ministry (Ministero degli Interni) and refer to 2011.

Regarding the output choice, as evident in the existing literature, it is difficult to find data that directly measure municipal production results: so, just surrogate measures of municipal demand are considered for performance indicators, often used as proxies for the related services provided to the citizens. In addition, there is no information about qualitative results of municipal activities: so, just quantitative data have been employed in the analysis. Moreover, the data available for some performance indicators sometimes have missing data with respect to some municipalities and certainly they become useless in the analysis. Taking into account these difficulties, the outputs proposed in the literature have been considered and function by function the variables have been selected.

The total population is considered as proxy for the various administrative tasks, as in Balaguer-Coll et al. (2007); Boetti et al. (2012); Da Cruz and Marques (2014); De Borger and Kerstens (1996); De Sousa and Stošić (2005); Geys et al. (2010); Kalb et al. (2012); Lo Storto (2016); Nakazawa

(2013). For the education services, the school age population (3-13 years old) has been taken as the catchment area of the services supplied by municipality (e.g., Boetti et al., 2012; Borge et al., 2008; De Borger and Kerstens, 1996; De Borger et al., 1994; Geys et al., 2010; Geys and Moesen, 2009; Kalb et al., 2012). Regarding social services, the output is given by the number of municipal citizens from 0 to 5 years old (for kindergarten and school canteen services) plus the number of the over 65 (for elderly provision) plus the number of immigrants (immigration needs), as in Afonso and Fernandes (2006); Boetti et al. (2012); Borge et al. (2008); De Borger and Kerstens (1996); Geys et al. (2010); Kalb et al. (2012); Nakazawa (2013). Local police activities as well as the road maintenance and local mobility function are measured by the total amount of kilometres of roads to be supervised/maintained and by the amount of resident population plus the average annual tourist presence since they are considered as proxy of the potential users of these services (e.g., Afonso and Venâncio, 2016; Boetti et al., 2012; Da Cruz and Marques, 2014; Geys and Moesen, 2009). In compliance with part of the existing literature, even the size of the municipal area could have been chosen as output at least for the administrative services, for the local police function and for the road maintenance one. Despite this, the geographical and socio-economic characteristics of Tuscan municipalities make this choice inappropriate. A preliminary statistical analysis has shown a strong bias in the outcomes due to the high heterogeneity related to this variable. However, given the importance of this aspect, the size of the municipal area is taken into account in the econometric analysis by means of the variable "density".

Data are collected from the statistical database DEMO ISTAT, the Mobility and Transport Regional Observatory and Tuscany Region survey. They all refer to 2011, consistently with the expenditure side, and they cover 282 Tuscan municipalities. Despite Tuscany has 287 municipalities, data were not available for two of them and three municipalities have been detected as outliers. In fact, from a first analysis on municipalities' features, Firenze has been considered absolutely out of scale in comparison with all the other municipalities. This intuition has been confirmed by a super-efficiency DEA analysis which has been performed to detect outliers (see for all Banker and Chang, 2006). Actually, a super-efficiency DEA model has been run function by function: two more municipalities have been detected as outliers and therefore dropped from the current analysis.

Table 4.1 and Table 4.2 present the dataset descriptive statistics respectively for the relevant input and output variables.

Table 4.1: Descriptive statistics for DEA dataset – Input.

INPUT	N	Mean	Stdev	Min	Max
General administration (10^3 €)	282	2843.708	4788.804	143.12	41457.83
Local police (10^3 €)	282	566.7216	1211.233	2.717	10696.06
Educational services (10^3 €)	282	1177.73	2146.233	41.306	18580.92
Road maintenance and local mobility (10^3 €)	282	857.3745	1937.075	15.823	18751.71
Social services (10^3 €)	282	1653.309	3730.047	5.042	35413.89

Table 4.2: Descriptive statistics for DEA dataset – Output.

OUTPUT	N	Mean	Stdev	Min	Max
Total population	282	11650.6	20235.62	394	184885
Length of roads (Km)	282	139280.9	157103.9	0	1353082
Population + Tourist presence	282	11998.32	20536.98	493.4521	186104
Population 3-13	282	1107.486	1954.895	20	19640
Population 0-5 + Over 65 + Immigrants	282	4348.447	7763.503	142	77943

4.4.2 Results

In this section, the DEA efficiency scores for 282 Tuscan municipalities are presented; they are computed by Coelli's software "DEAP Version 2.1: A Data Envelopment Analysis (Computer Program)" (Coelli, 1996). The assessment of expenditure performance is expressed in terms of DEA scores by values between 0 and 1: the municipalities with a score equal to one are those that are fully efficient.

As explained in the previous section, for each fundamental municipal function a VRS analysis is done and for the global analysis the DEA-like composite indicator is computed. Table 4.3 presents in a synthetic way the main descriptive statistics of the results both for each municipal area and for the overall level: the full list of the results is in Appendix C.1.

Table 4.3: Descriptive statistics of the efficiency scores at local and global level.

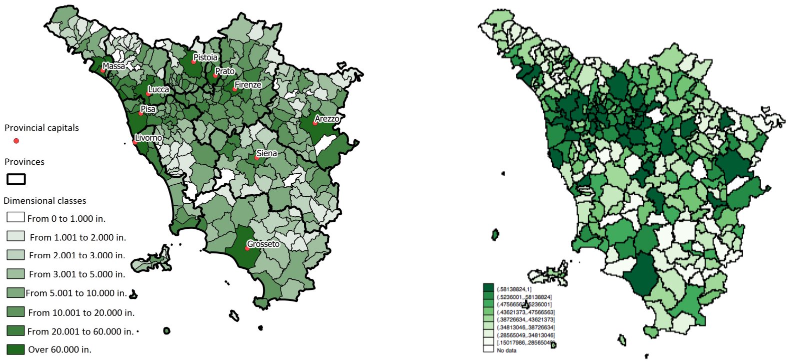
	Mean	Stdev	Min	Max	Percentiles				
					10°	25°	50°	75°	90°
General administration	0.59	0.19	0.15	1.00	0.37	0.46	0.58	0.73	0.84
Local police	0.43	0.23	0.02	1.00	0.14	0.26	0.42	0.56	0.76
Educational services	0.43	0.16	0.13	1.00	0.26	0.32	0.40	0.51	0.61
Road maintenance	0.35	0.20	0.03	1.00	0.14	0.20	0.30	0.43	0.61
Social services	0.45	0.24	0.02	1.00	0.18	0.27	0.40	0.57	0.83
DEA-like Overall	0.44	0.13	0.15	1.00	0.27	0.35	0.44	0.52	0.61

A question may naturally arise: do specific municipal features affect local public expenditure management and its efficiency? To address this question, four main municipal characteristics are considered in line with the Tuscan hallmarks: size, geography, tourism degree and socio-economic structure through the local labour system classification. Since from a qualitative point of view the outcomes are the same both at non-aggregate and at overall level, for the sake of brevity only the descriptive statistics for the composite indicator are reported below in Table 4.4 and 4.5, even if it should be taken in mind that the efficient municipalities considered as peer for all the other inefficient ones vary according to each function, both in terms of number of efficient units and in terms of municipal typology. In the following, the four listed aspects are shortly analysed.

To explore the effect of municipal size, the efficiency scores are clustered in eight different classes: as the municipal population size increases, the average of the efficiency scores among each class increases. As regards the biggest class, there is the highest minimum value of efficiency score and the highest maximum value, equal to one, meaning that according to this analysis the most and fully efficient municipality belongs to this class and it is the provincial capital Prato. Broadly speaking, all the provincial capitals result to be the most efficient municipalities: this can be seen graphically from Figure 4.1 [3]. Moreover, the geographical distribution of the municipal population size reflects a similar distribution in terms of the expenditure efficiency scores: the darker the area,

[3] Figure 4.1(b) is obtained by "Stata" program.

the more populated and efficient the municipality. This evidence makes stronger the reasoning about municipal size: the bigger is the municipal catchment area, the lower the average cost in the provision of municipal services, which in turn makes possible to provide more differentiated and complex services. In particular, as regards the small municipalities, the inefficiency source can be related to the presence of too many small fragmented municipalities: this might suggest an aggregation among the smallest municipalities to exploit scale economies, in line with the legislative measures proposed by the Tuscany region to overcome this problematic aspect.



(a) By dimensional classes (b) By global efficiency scores

Figure 4.1: Geographical distribution comparison.

To check whether the municipal geography plays a role in the resources management, the "mountain" feature is considered. In line with the Italian legislation (Law 991/52, Law 657/57 and Law 142/90), the "mountain classification" distinguishes three categories: totally mountain, partially mountain and non-mountain. The lowest efficiency scores are present in the mountain municipalities: certainly, the difficult terrain and the smallest presence of the resident population make more inefficient the provision of the services.

As tourism is a very important Tuscan feature, the degree of municipal tourism involvement is obtained dividing into quartiles the ordered per capita tourist presence. The highest level of inefficiency is present in the municipalities with high level of tourism, while the opposite holds for those municipalities with very low level of tourism. In general, it's possible to observe that considering an increasing level of tourism, the average level of efficiency systematically decreases. Certainly, a remark must be made when considering the tourist presence: vacation property owners are not taken into account, even though they might represent a non-negligible part of the catchment area of the municipal services and so lower the inefficiency scores. Anyhow, especially the tourist municipalities subject to strong seasonality face higher costs than others (e.g. this is the case of the sea places).

Finally, local labour systems are used to investigate the Tuscan municipalities' socio-economic structure: they are territorial units of daily activities of the population that lives and works there and consist of several adjacent municipalities, geographically and statistically comparable with each other. The classification is based on the ISTAT (Italian National Institute of Statistics) elaboration. The lowest average efficiency level is present in the systems without specialization, while the opposite holds for the urban systems and the manufacturing systems in textile, leather and clothing. In relation to this last mentioned class, it's worth noting that the obtained most efficient municipality, Prato, belongs precisely to it.

Table 4.4: Descriptive statistics of global efficiency scores.

	Mean	Stdev	Min	Max
Size class				
From 0 to 1.000 inhab.	0.37	0.12	0.16	0.58
From 1.001 to 2.000 inhab.	0.32	0.09	0.20	0.62
From 2.001 to 3.000 inhab.	0.36	0.06	0.19	0.51
From 3.001 to 5.000 inhab.	0.39	0.08	0.22	0.57
From 5.001 to 10.000 inhab.	0.45	0.10	0.15	0.71
From 10.001 to 20.000 inhab.	0.54	0.09	0.37	0.75
From 20.001 to 60.000 inhab.	0.55	0.12	0.35	0.82
Over 60.000 inhab.	0.71	0.16	0.44	1.00
Mountain class				
Non-mountain	0.49	0.13	0.15	1.00
Partially mountain	0.49	0.14	0.22	0.92
Totally mountain	0.37	0.10	0.16	0.61

Table 4.5: Descriptive statistics of global efficiency scores (cont'd).

	Mean	Stdev	Min	Max
Tourism class				
Very low tourism	0.48	0.14	0.20	0.82
Low tourism	0.47	0.15	0.23	1.00
Medium tourism	0.45	0.12	0.22	0.80
High tourism	0.38	0.11	0.15	0.62
Local labour system class				
Without specialization	0.37	0.08	0.24	0.57
Urban systems	0.49	0.13	0.21	0.80
Tourism and agricultural vocation	0.39	0.14	0.19	0.82
Manufacturing in the textile, leather and clothing	0.49	0.14	0.20	1.00
Other manufacturing made in Italy	0.45	0.14	0.16	0.92
Heavy manufacturing	0.41	0.11	0.15	0.66

4.4.3 A different way of aggregation

As pointed out before, there are several ways to aggregate the fundamental function efficiency scores into a composite indicator, according to the different assigned system of weights. Therefore, at this point of the analysis, a different composite indicator is proposed: it is computed as the weighted average of the function efficiency scores, using the weights they have in the total expenditure, so to represent in a synthetic way the average municipal spending efficiency results.

First of all, it is a useful tool as a "robustness" check of the results obtained with the DEA-like aggregating approach: Figure 4.2-4.5 show that for the four mentioned municipal features the comparison of the two indicators exhibits the same trend. Intuitively, this might suggest two considerations. First of all, the DEA-like CI has been constructed so to give higher importance to the functions that indeed represent a higher share of expenditure in the budget allocation. Moreover, municipalities devote the more attention in the management of the resources the greater is the relevance of the function (as by construction higher weights are assigned to higher non-aggregate efficiency scores). This evidence might depend on the structure of the data or be generalized as a property of the model: this question represents an interesting topic for further research.

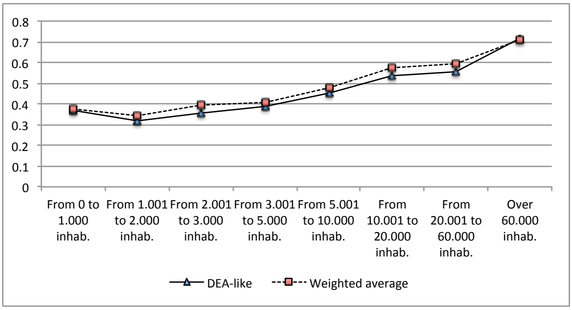


Figure 4.2: Comparison between composite indicators scores by dimensional classes

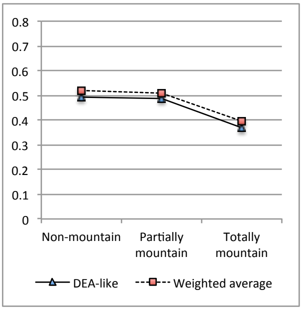


Figure 4.3: Comparison between composite indicators scores by mountain classes

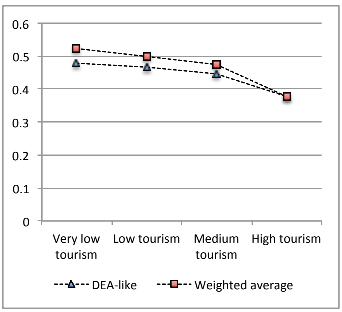


Figure 4.4: Comparison between composite indicators scores by tourism classes

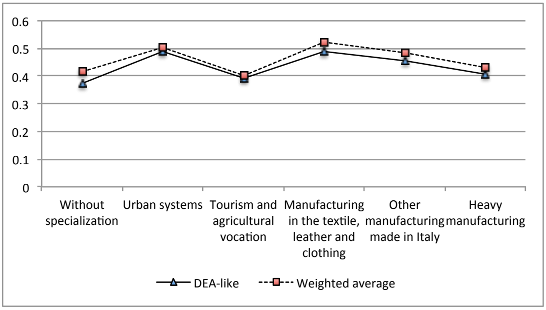


Figure 4.5: Comparison between composite indicators scores by local labour system classes

Furthermore, though from a technical point of view the proposal of this kind of indicator might be simpler, from a normative point of view it might give more intuitive and then useful suggestions for the legislator and for the local governments to find room of improvements in resource management. To reach this goal, both the composite indicator obtained by using each municipal expenditure composition and the one obtained by considering the Tuscan mean expenditure composition are necessary: the descriptive statistics are listed in Table 4.6. Using this approach, on the one hand, it becomes possible to make some considerations about the effect of the municipal expenditure allocation among the different functions on the average inefficiency. On the other hand, the units under analysis can be divided into groups according to different level of efficiency, the most frequent municipal features can be identified and compared with the commented evidences of the previous section.

Table 4.6: Descriptive statistics of the scores for municipal and Tuscan average weights.

	Mean	Stdev	Min	Max	Percentiles				
					10°	25°	50°	75°	90°
Municipal weights	0.47	0.14	0.13	1.00	0.30	0.37	0.46	0.57	0.64
Tuscan weights	0.49	0.14	0.16	1.00	0.33	0.39	0.47	0.59	0.66

The proposed procedure is the following. First of all, the municipalities have to be divided according to two features. The first regards the relative level of efficiency: it is the difference between each municipal composite indicator (computed considering the own municipal expenditure composition) and the median of all these computed indicators, so to have the relative efficiency of each municipality (to be noticed that the median quite coincides with the mean). The second feature regards the expenditure composition: it is the difference between each municipal composite indicator computed considering the municipal expenditure composition and the one obtained taking the Tuscan average expenditure composition. If this difference is positive, this suggests that the municipality has chosen a composition that allows it to achieve a better level of average efficiency rather than in any other resources allocation; if the difference is negative, then the municipality has chosen a composition that brings it to achieve a worse level of efficiency.

Then, Figure 4.6 shows the combination of these two dimensions in a graphical and intuitive way to distinguish four groups of municipalities: on the vertical axis there is the relative efficiency, while on the horizontal axis the expenditure composition aspect is considered. Municipalities are laid out into four quadrants according to the following way: for y positive values the municipalities belong to the Efficient quadrants, that is they are more efficient than the median, while for x positive values the municipalities belong to the Better quadrants, that is they have an expenditure composition that allows them to achieve a better level of average efficiency; the opposite reasoning holds respectively for the Inefficient and the Worse quadrants.

As evident, it can be said that the municipalities in the Efficient-Worse and Inefficient-Worse quadrant have possible room of improvement in the efficiency level just changing a little the composition of the expenditure. Certainly, this suggestion should be handled carefully, especially for two reasons: the change in expenditure brings a change in the DEA model input, so to modify endogenously the level of the efficiency; secondly, especially for the smallest municipalities there are some binding thresholds of expenditure that cannot be avoided. Even so, the municipalities in the Inefficient-Worse and Inefficient-Better quadrant certainly could improve their level of efficiency at least solving the present mismanagement problems and their causes. So, in conclusion, the Efficient-Better quadrant seems to collect the municipalities that behave better, according to this analysis.

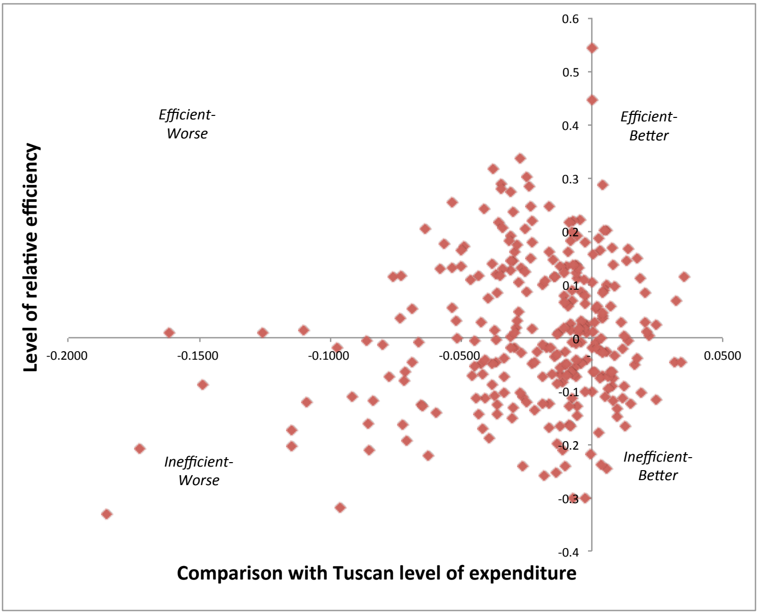


Figure 4.6: Municipalities by relative efficiency and expenditure composition.

In a synthetic way, Table 4.7 shows the main features of each quadrant according to the already used dimensional, mountain, tourism and local labour system classes and referring the number of present municipalities (in the table shortly DMUs). The dimensional class which reveals the highest percentage of the best performing municipalities (Efficient-Better quadrant) is the one with over sixty thousands of inhabitants. Referring to the other classifications, the highest percentage of the best performing municipalities lies in the non-mountain class, in the low-tourism one and in the manufacturing systems in the textile, leather and clothing one. As evident, these features recall those already presented in the previous results. Similar way of reasoning can be given for the other quadrants.

Table 4.7: Descriptive statistics of each quadrant.

	Efficient-Better quadrant		Efficient-Worse quadrant		Inefficient-Better quadrant		Inefficient-Worse quadrant		
Dimensional class	DMUs	%	DMUs	%	DMUs	%	DMUs	%	TOTAL
From 0 to 1.000 inhab.	3	18%	2	12%	1	6%	11	65%	17
From 1.001 to 2.000 inhab.	1	3%	3	8%	12	30%	24	60%	40
From 2.001 to 3.000 inhab.	3	11%	3	11%	8	29%	14	50%	28
From 3.001 to 5.000 inhab.	6	13%	9	19%	7	15%	26	54%	48
From 5.001 to 10.000 inhab.	8	13%	29	46%	6	10%	20	32%	63
From 10.001 to 20.000 inhab.	9	18%	34	69%	3	6%	3	6%	49
From 20.001 to 60.000 inhab.	4	15%	18	67%	1	4%	4	15%	27
Over 60.000 inhab.	2	20%	7	70%	0	0%	1	10%	10
TOTAL	36	13%	105	37%	38	13%	103	37%	282
Mountain class	DMUs	%	DMUs	%	DMUs	%	DMUs	%	TOTAL
Non-mountain	21	16%	61	48%	17	13%	29	23%	128
Partially mountain	5	12%	23	55%	2	5%	12	29%	42
Totally mountain	10	9%	21	19%	19	17%	62	55%	112
TOTAL	36	13%	105	37%	38	13%	103	37%	282
Tourism class	DMUs	%	DMUs	%	DMUs	%	DMUs	%	TOTAL
Very low tourism	15	21%	33	47%	3	4%	19	27%	70
Low tourism	8	11%	31	44%	11	15%	21	30%	71
Medium tourism	7	10%	30	43%	8	11%	25	36%	70
High tourism	6	8%	11	15%	16	23%	38	54%	71
TOTAL	36	13%	105	37%	38	13%	103	37%	282
Local labour system class	DMUs	%	DMUs	%	DMUs	%	DMUs	%	TOTAL
Systems without specialization	4	12%	8	24%	3	9%	18	55%	33
Urban systems	3	7%	22	51%	4	9%	14	33%	43
Tourism and agricultural vocation systems	5	12%	9	21%	9	21%	20	47%	43
Manufacturing systems in the textile, leather and clothing	12	16%	37	50%	7	9%	18	24%	74
Other manufacturing systems made in Italy	5	13%	18	45%	5	13%	12	30%	40
Heavy manufacturing systems	7	14%	11	22%	10	20%	21	43%	49
TOTAL	36	13%	105	37%	38	13%	103	37%	282

4.4.4 Efficiency explanatory variables: Tobit regression

To summarize the evidence of the Sections 4.4.2 and 4.4.3 and in compliance with some of the existing literature (e.g. in Athanassopoulos and Triantis, 1998; Boetti et al., 2012; De Borger and Kerstens, 1996; Worthington, 2000), a Tobit regression is implemented to explain the global efficiency scores. In the literature, the consistency and the validity of the commonly used Tobit and OLS regression models have been put under scrutiny. To address these issues, additional techniques have been proposed, as for example sensitivity analysis and bootstrap procedures (see e.g. Da Cruz and Marques (2014); Lo Storto (2016)): the last mentioned is the one considered in this analysis. It's worth pointing out that also other estimation models such as the Maximum Likelihood Estimation-MLE (see e.g. Poveda, 2012) and the Generalised Method of Moment-GMM estimation (see e.g. Poveda, 2014) have been recently used to detect the effect of environmental variables on the efficiency scores. Overall, in this context the main interest is to get in a synthetic way what are the underlying causes of the estimated efficiency gaps and the potential determinants of municipal inefficiency, summarizing the main outcomes of the analysis so far proposed.

The explanatory variables are chosen considering the existing literature and refer to three types of variables related to: i) the economic and financial aspects ii) the municipal characteristics and iii) the politics components. The two variables, "autonomy" and "revenues" belong to the first type. As in Boetti et al. (2012), the variable "autonomy" measures the degree of accountability of local governments with respect to citizens, here defined as the ratio of local taxes to the total expenditure. The variable "revenues", given by the ratio of total revenues and total resident population, is a proxy of soft budget constraints (Reingewertz, 2012): governments with less revenues are assumed to be more careful to control their expenditure (see also Kornai et al., 2003; Šťastná and Gregor, 2015). The municipal characteristics refer to geo-demographic and touristic aspects that could explain efficiency scores. Demographical aspects are measured through dummy variables (e.g. Boetti et al., 2012; Doumpos and Cohen, 2014; Šťastná and Gregor, 2015): "Dim1" for municipalities from 0 to 5.000 inhabitants, "Dim2" for municipalities from 5.000 to 10.000 inhabitants, "Dim3" for municipalities from 10.000 to 20.000 inhabitants, "Dim4" for municipalities from 20.000 to 60.000 inhabitants and "Dim5" for municipalities over 60.000 inhabitants. Regarding geo-demographic aspects, as in Athanassopoulos and Triantis

(1998); Doumpos and Cohen (2014); Geys et al. (2010); Kalb et al. (2012); Lo Storto (2016), the "density" of the municipality and the variable "mountain", equal to 1 for mountain municipalities according to the Italian legislation, are considered. According to Balaguer-Coll et al. (2013); Benito et al. (2010); Cuadrado-Ballesteros et al. (2013); Da Cruz and Marques (2014); De Sousa and Stošić (2005), the importance of tourism is taken into account through the variable "tourism", defined as the ratio between the average annual tourist presence and the total population. The variable "second mandate" is finally introduced to take into account the potential influence of political factors on efficiency scores. Data are collected from the municipal balance sheets, the statistical databases DEMO ISTAT and ISTAT, Tuscany Region survey and ANCI TOSCANA and they all refer to 2011.

The Tobit regression is run for both the bias-corrected[4] global efficiency scores, the DEA-like and the average one, considered as the dependent variable, and implemented by the software "Stata". Table 4.8 contains the Tobit results for both the bias-corrected global efficiency scores: if an explanatory variable has a positive sign, it positively affects the efficiency and if it has a negative sign, the opposite holds.

The results are very similar for both the model specifications, statistically significant and in line with the literature's main findings. The economic and financial variables have the expected sign. In fact, the degree of accountability measured by the variable "autonomy" positively affects the efficiency score, while the budget constraints measured by the variable "revenues" have a negative impact on it. The effect of the municipal characteristics on the efficiency scores confirms the descriptive analysis of the two previous sections. As can be seen from the coefficients of the dimensional dummy variables, when municipal size increases, the efficiency score increases as well. So, consistently with the in-depth analysis stemming from both the efficiency composite indicators adoption, the larger the municipalities the more efficient the expenditure management. Additionally, the results of Wald tests show that the hypothesis of

[4]The bias-corrected efficiency scores are computed using 1000 replicates in the first stage. Some of the current literature suggests further bias-corrections to get more robust results (see e.g. Vidoli and Mazziotta, 2013). For the present analysis, several Robust CIs have been computed: as they have a strong correlation with the CI obtained with the bias-corrected efficiency scores, the CI adopted in the analysis is considered sufficiently robust. The computations have been performed by using the statistical software R (Team, 2015), in particular the packages "Benchmarking" (Bogetoft and Otto, 2010) and "Compind" (Vidoli et al., 2015). Results are available upon request. The authors are grateful to an anonymous referee for suggesting these additional robustness checks.

Table 4.8: Tobit results.

VARIABLE	DEA-like	Average
AUTONOMY	0.12425***	0.12842***
REVENUES	-0.00006***	-0.00005***
DIM2	0.06775***	0.04688***
DIM3	0.13973***	0.10775***
DIM4	0.11424***	0.09983***
DIM5	0.17529***	0.12472***
DENSITY	0.00004	0.00002
MOUNTAIN	-0.05683***	-0.04220***
TOURISM	-0.35370***	-0.31774***
SECOND MANDATE	0.02884*	0.02066*
CONSTANT	0.54073***	0.35756***

* 5% significance, ** 1% significance, *** 0.1% significance

no differences between Dim2, Dim3, Dim4 and Dim5 is strongly rejected. The variable "density" shows a positive, though not statistically significant, impact on the efficiency index (as in the majority of the studies). The negative impact of the variable "mountain" could be linked to the demographical aspects, as the mountain municipalities tend to be of a smaller size. Furthermore, the coefficient of the variable "tourism" confirms what shown in the descriptive analysis of the previous sections: municipalities with high level of tourism tend to be less efficient. Finally, the political variable "second mandate" has a positive coefficient: the incumbent politicians, in an effort to signal their competence to the voters so as to increase their chances to be reelected, tend to enlarge spending (inefficiently) when they are close to new elections (Rogoff and Sibert, 1988). Administrations at the second mandate don't have the possibility to be elected again and therefore the positive effect on the spending efficiency is in line with the economic theory (see also Boetti et al., 2012).

4.5 Conclusions

In this chapter, the efficiency of Tuscan municipal expenditure is under scrutiny by means of Data Envelopment Analysis. The data on municipal expenditure are taken from the available municipal balance sheets and the following functions are considered, given their importance on the total current expenditure: “General administration”, “Educational services”, “Social services”, “Road maintenance and local mobility” and “Local police”.

For the function by function analysis, a separate DEA model is run for each “fundamental” municipal area. For the overall analysis, addressing some methodological issues to compute the global efficiency score, the use of a DEA-like composite indicator is introduced, for the first time in this strand of literature. Moreover, a further composite indicator is proposed following another approach, strictly related to the municipal expenditure compositions. In this latter case, the efficiency score of each function enters in the global indicator with the same proportion that the given function has with respect to total expenditure. Although the two composite indicators are derived following two very different approaches, the conclusions are basically the same. The first approach is closer to the principle of DEA; the second one is more operative and it could suggest some normative indication to the policy-makers in terms of the expenditure distribution. In this light, the composite indicator obtained by the municipal weight is also compared with the indicator obtained by the Tuscan mean weight; there are possible suggestions as room for improvement for the inefficient municipalities: in some cases, just a change in the composition of the expenditure could bring an increase of the composite indicator efficiency score.

The results obtained through a DEA analysis and validated by the Tobit regression appear consistent and could be a starting point for the reallocation of the inefficient municipalities’ expenditure. In particular, some evidence about the long debated issue of the municipal size comes out. In fact, according to this analysis, the municipal size really affects the efficiency of public expenditure: the bigger is a municipality, the greater is its level of public spending efficiency, so that the regional measures to reduce the present fragmentation of the Tuscan territory seem to be in line with this evidence.

The performed analysis offers further insights from both a methodological point of view and an empirical one. Some of them are driven by the limitations of the analysis itself. It is well-know that a complete efficiency evaluation of local government activities should also include the quality of the services and citizens’ satisfaction. As the present chapter considers only quantitative data, a further stream of research could be the definition of suitable variables embedding both these qualitative aspects. Moreover, the way of combining qualitative and quantitative elements is still an open issue that should be investigated. The aggregation of these kinds of different data might suggest the necessity of the construction of other new composite indicators. With this regard, there exists a growing interest in the current scientific debate on the definition and the use of new composite indicators. Furthermore, the present study considers data referred just to one year. However, monitoring the changes on how public resources are spent over a larger period of time can represent a key point in the municipal spending efficiency analysis. For this reason, longitudinal data could be used to perform an intertemporal efficiency analysis by means of DEA window approaches.

In this chapter, we question whether the public expenditure is allocated properly across the local functions and accordingly whether it is managed in an efficient way. However, in the public spending analysis, also exploring how successfully resources are used in reaching the set goals is of primary concern: Chapter 5 investigates the resource effectiveness in a specific public area, the education sector.

Chapter 5

School Infrastructure Spending and Educational Outcomes in Northern Italy

5.1 Introduction

The goal of the policy-makers is both "doing things right", but also "doing right things". Chapter 4 has addressed the first objective, while the present chapter considers the second one, namely the evaluation of resource management in terms of effectiveness. In particular, public spending in the education sector is under analysis, as it represents one of the most important expenditure items in public spending, despite the fact that its effect on students' achievement is still under discussion.

Whether or not school spending has an impact on student outcomes is a highly debated issue in economics (Card and Krueger, 1996). The contemporary literature has been pioneered by Coleman (1966) in a prominent report published by the US Government in 1966, whose main conclusion is that school funding does not play a central role in determining students' achievement. A wealth of studies follow in the footsteps of Coleman (1966) and explore the relation between resources and educational outcomes (see e.g., Jackson et al., 2016; Neilson and Zimmerman, 2014). Overall, there is a lack of agreement on the impact of funding on students' performance. Whereas in a meta-analysis Greenwald et al.

(1996, p.384) conclude that "school resources are systematically related to student achievement and that these relations are large enough to be educationally important," many subsequent studies find little or no effect (see e.g., Card and Krueger, 1996; Hanushek, 1996).

This disagreement is perhaps not very surprising as most of these studies face severe difficulties in attempting to unravel a causal relationship between school spending and educational outcome. Counterfactual outcomes are sensitive to the choice of the estimator and the identification strategy to address the endogeneity of school resources. Although previous studies have made a good deal of progress in dealing with the joint determination of educational inputs and outputs, modest estimated effects of school spending could be a consequence of unresolved endogeneity biases (see Jackson et al., 2016). At the same time, studies often explore very heterogeneous inputs of the educational production process. Jones and Zimmer (2001) note that most of the literature focuses on school-specific inputs, school organization inputs (e.g., class size), environmental characteristics and socioeconomic (family) characteristics, but neglects capital inputs such as school infrastructure. In fact, there are only a handful of studies on the school infrastructure-students' learning relationship and they focus predominantly on the US School System. Aaronson and Mazumder (2011) investigate the impact of the so-called "Rosenwald initiative" in the US between 1914 and 1931 and find that substantial improvements to school quality and access in relatively deprived environments are followed by large productivity gains. Neilson and Zimmerman (2014) find strong evidence that school construction programs led, among other outcomes, to sustained gains in reading scores for elementary and middle school students. Yet, Cellini et al. (2010) and Martorell et al. (2016), who focus more specifically on school facility investments, find little evidence that spending on facilities generates improvements in student achievement.

Against this background, we explore whether spending on physical infrastructure affects student outcomes by focusing on test scores in mathematics and Italian language using data on Italian state high schools. The issue of school capital funding features prominently in the public debate, and in many countries the lack of investment remains a pressing priority for state schools, where many school principals believe that schools are not "fit for purpose" (Guardian, 27/01/2015).[1] In Italy,

[1] To give an example, in 2017, the Australian government will bring forward $200 million in capital investment to fast track state school infrastructure throughout Queens-

school principals have long lamented that poorly maintained school facilities and a lack of funding to conduct essential repairs prevent schools from delivering effective schooling (Corriere della Sera, 18/07/2017). This is in line with theoretical arguments put forward by educational researchers, social psychologists and sociologists on the importance of the physical environment of schools and the condition of their facilities in explaining variation in students' learning across schools (Bakó-Biró et al., 2012; Earthman, 2002; Haverinen-Shaughnessy et al., 2015; Mendell and Heath, 2004). Specifically, this literature has stressed the role of social norms, conformity and social signalling in the school environment (Branham, 2004): the so-called "*broken windows theory*" (Wilson and Kelling, 1982) relies on the premise that the school environment "communicates" to students and that "good signals" correlate with a more efficient learning process. The main idea is that a well maintained school environment helps to create an atmosphere of order and comfortable place to study. Lawrence (2003) reviews a number of studies exploring how the condition of the school facility affects the health and morale of staff. This interpretation may help to clarify the apparently conflicting results seen in the literature so far and identifies a potential pathway to explain the direction of educational outcome's change in response to infrastructure spending. An adequate school environment is perceived as the foundation for building the future: "the school building should be telling a country's potential" and safety and infrastructure quality should be the basis of a good school" (Repubblica, 16/10/2017). On the one hand, a safe and clean school environment provides important signals to students that the school is well managed, that teachers enforce discipline in the classroom, and that e.g., bully behaviour is not tolerated. On the other hand, unhealthy and unsafe buildings, with e.g., broken windows, graffiti, nonfunctioning toilets, poor lighting, inoperative heating and cooling systems, leaking roofs, signal a lack of attention and respect for the students, who either put less efforts or distract colleagues and disrupt the learning environment, as they perceive lower costs and risks of detection. Students in well-maintained schools are therefore more likely to focus on academic challenges than those who are distracted or depressed

land (`https://goo.gl/GGe1Pf`). In 2015-16, the UK Department for Education spent GBP 4.5 billion in capital funding, and the National Audit Office has predicted that it will take a further GBP 6.7 billion investment to bring all schools up to scratch (`https://goo.gl/SQzHDE`). In Germany, Martin Schulz, leader of the Social Democrats, vowed to pour billions into crumbling schools infrastructure in campaigning for 2017 September's election (see FT, 17/07/2017).

by the poorly maintained facilitates, avoiding chronic distractions and missed school days. In addition, low-quality facilities undermine effort among students and especially among the low-achieving ones. By the same token, physical conditions also affect teachers' feelings of effectiveness and sense of personal safety in the classrooms. Moreover, badly maintained buildings might not have the infrastructure to support the latest technology or could lack modernized labs for science education (Martorell et al., 2016). Finally, to be sure that there are no other confounding mechanisms, we specify what the extra-funding has been used for: we gathered this information available upon request to the reference province or to any authority in charge of the reparation implementation. In our sample, interventions to the schoolhouses have been not invasive and funds have been mostly used for the painting of scratched wall, the lightening of a gym and the fixing of the heating system for example: money has been used to support the school infrastructure and not to buy PCs or other technological devices.

To handle the endogeneity of idiosyncratic changes in school funding, we use two strategies. First, we employ a quasi-experimental design and make use of information on the extra funding that a specific group of schools received in the aftermath of the 2012 Northern Italy earthquake. In May 2012, the seismic events in Northern Italy caused considerable damage to state buildings and prompted specific interventions for the mitigation of the seismic risk. As a result, a large number of undamaged schools, but close enough to the areas affected by the earthquake, received large extra funds to modernize and improve the quality of their buildings as well as to mitigate their vulnerability to earthquakes. We compute the differential effect of receiving extra funds on the treatment group, i.e., undamaged schools outside the earthquake area, that were awarded special funding, versus a control group of schools in neighbouring municipalities. The schools in the control group are in areas sufficiently far from the earthquake epicenter and at low risk of future seismic activities; therefore these schools are both undamaged as well as unfunded. This strategy allows us to estimate whether being a recipient of funding increases students' achievement. Second, to evaluate the elasticity of test scores with respect to funding, we implement an instrumental variables (IV) identification strategy. In particular, we use seismic hazard maps and exploit exogenous values of peak ground acceleration (henceforth PGA), which explains much of the variation in the amount of funds received. Taken together, our results suggest that improving the quality of school buildings has a positive effect on students' achievements. Moreover, we find that low-achieving students benefit the most from improved physical infrastructure.

The rest of the chapter proceeds as follows. We give a sketch of the Italian school system in Section 5.2 and we describe the quake-related events and the policies implemented afterwards in Section 5.3, so to put into context our empirical analysis. Section 5.4 describes the data, while we lay out our identification strategy in Section 5.5. Section 5.6 provides a discussion of our main results. Finally, we conclude in Section 5.7.

5.2 The Italian school system framework

The work presented in this chapter aims at detecting the effect of school infrastructure spending on educational achievement. To give an overview of the institutional setting our analysis is built on, in this section we describe the schooling context in Italy, with a particular focus on school funding and student outcome evaluation aspects.

Italian schooling is compulsory from ages 6 to 16. Grades are grouped into three stages: primary school covers the first five grades; lower secondary school gathers grades from sixth to eight; high school runs from the ninth to the thirteenth grade. The last two grades are not compulsory, but the withdrawal rate is quite small.[2]

The Italian school system is predominantly public and its funding is mainly managed by the Ministry of Education. Schools' funding varies across the three stages, but generally not within them. Schools are by far mostly funded by public funds. The main source, according to which funds are split across schools, is the "*Fondo per il funzionamento del sistema scolastico*". The criteria that regulate its allocation are basically two, as stated in the ministerial decree # 21/2007:

(i) A fixed component that takes into account the school type. Lower grades (primary and lower secondary) schools receive less funds than high schools. Specifically, the first group of schools receives an amount equal to €1.100; the second one €2.000.

(ii) A component that varies according to the number of enrolled students. However, the amount of euros that each school receives per student varies according to its type too: primary and lower secondary school receive 8 euros per student; as concerns the high

[2]Source: ISTAT http://www.istat.it/it/archivio/17290.

schools, the state contribution per student ranges from 12 euros for the *licei* to 48 euros for technical high schools.

Other transfers from the state to the school are made to cover schools' additional costs such as short term teacher vacancies or to buy goods and services to be used to assist students with disabilities. Teachers are mainly paid by the state according to their seniority. No regard is made on their qualifications, performance or conduct. The projects that could be awarded either at the state-level or at the European Union (EU)-level are the last source of funding. The likelihood of winning the project is however uncertain, so their distribution across the Italian schools is quite rare. Therefore, the basic structure of the Italian system is one in which schools are funded through public money and such funds are mainly managed by the Ministry of Education.

Figure 5.2 shows the financial flow diagram as depicted by the European Commission (2014), to give an overview of the transfer of public resources awarded in cash or in kind to schools. [3] Specifically, as concerns capital goods, transfers in cash are gathered from the ministry of education, of economy and of interior and allocated across the regions, which in turn give them to the provincial councils. The same procedure applies for the kind of funds assigned after the earthquake we deal with in our analysis, as we further explain in the next section.

From the discussion made above, it clearly follows that there are no measures that possibly relate to students' achievements among the criteria used to allocate funds across schools. Moreover, up to few years ago the only index considered meaningful to signal the school "quality" was its size, that is the number of students enrolled in the school. Only since 2008 this proxy has been complemented by an assessment of Italian students' skills and achievement carried out by an independent public agency, namely the National Institute for the Educational Evaluation of Instruction and Training (known by the Italian acronym INVALSI). Specifically, from the school year 2010/11, at the end of each year INVALSI administers a standardized test of Mathematics and Italian language skills to students in second, fifth, eight and tenth grade according to the following schedule: on May 3 the second and the fifth graders take the Italian language test; then the same students take the Mathe-

[3] In Italy, regions and provincial councils are mainly in charge of allocating school funds at high school level. Instead, municipalities are mainly in charge at primary and lower secondary school level, together with kindergarten. For a more in-depth discussion, European Commission (2014).

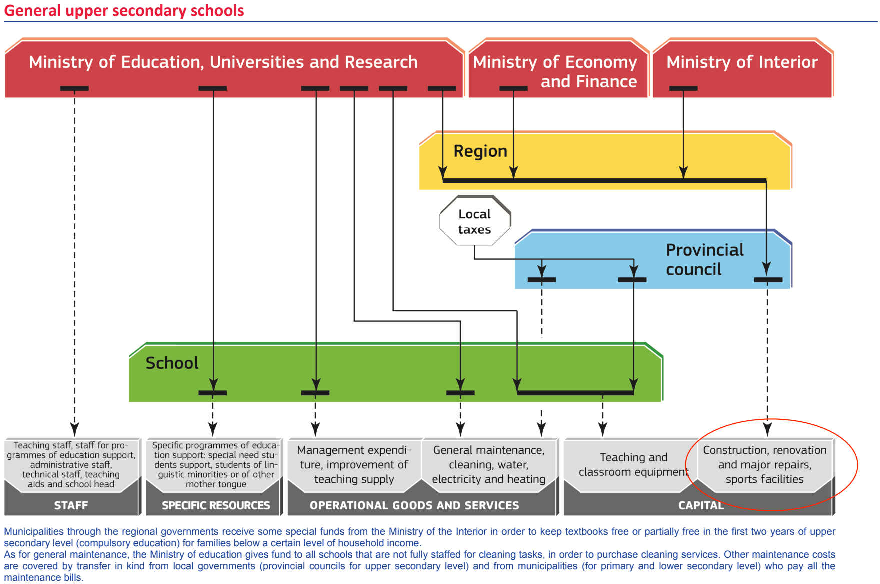


Figure 5.1: Financial flow diagram for the general upper secondary schools
Source: European Commission (2014)

matics test. Eighth graders take the two tests on the same day, on June 15, whereas the tenth graders take both them the 9th of May. The great advantage of such procedure is that those tests are administered at the same time to all students of the same grade and, above all, they are standardized, so eligible to make comparison across students from different classes and schools.

Despite the fact that the INVALSI test results are publicly presented once a year (at the end of the school year), highly debated by the national broadcasters and generally discussed in each school in September (at the beginning of the new school year), they have no impact whatsoever on students' life. In fact, they do not contribute to the final mark assigned to students[4] and, most importantly, they do not affect the way public resources are allocated across schools: accordingly, school funding im-

[4]The only exception is represented by the test administrated during the national exam students take to transit from the middle school to the high school (eight grade students). However, its contribution amounts only to 2% of the final mark.

plications are totally absent.

5.3 The Northern Italy earthquake overview

To investigate the effect of school capital spending on student outcome, we use information on the extra funding that a group of schools received after a natural disaster, namely the 2012 Northern Italy earthquake. In the following, we give a brief sketch of the seismic-related events and the post-quake interventions so to put into context the two intertwined empirical strategies proposed for our causal estimation as described in Section 5.5.

5.3.1 The May 2012 earthquake

On May 20, 2012, an earthquake of magnitude 5.9 ML hit a wide portion of the Po Valley in the Northern part of Italy. The epicenter was located near the town of Finale Emilia (MO), about 30 km west of the city of Ferrara, and the earthquake involved exclusively an area of about 3.5 thousand square kilometers across the three regions of Emilia Romagna, Veneto and Lombardy. The provinces affected by the earthquake were those of Ferrara, Modena, Mantua, Bologna, Reggio Emilia, and Rovigo, as officially stated in the law # 122/2012.

Figure 5.3.1 shows the epicentral area affected by the seism. Modena, Bologna, and Ferrara were the most affected provinces. In the first province nearly thousand square kilometers were damaged by the earthquake (about 36% of its territory). In the province of Bologna the area involved was 930 square kilometers (the 25% of the total). Finally, 31% of the province of Ferrara reported damages, in a territory of 818 square kilometers. The other three provinces were marginally impacted, with a total hit territory amounting to a thousand square kilometers.

5.3.2 The damage evaluation

After the series of seismic events, a macroseismic survey has been performed by the National Institute of Geophysics and Volcanology (known by the Italian acronym INGV) with the aim of assessing the amount of physical damages in the area. The survey assigns an intensity value proportional to the percentage of the damaged buildings in each locality—considered jointly with their proper vulnerability and damage level. For

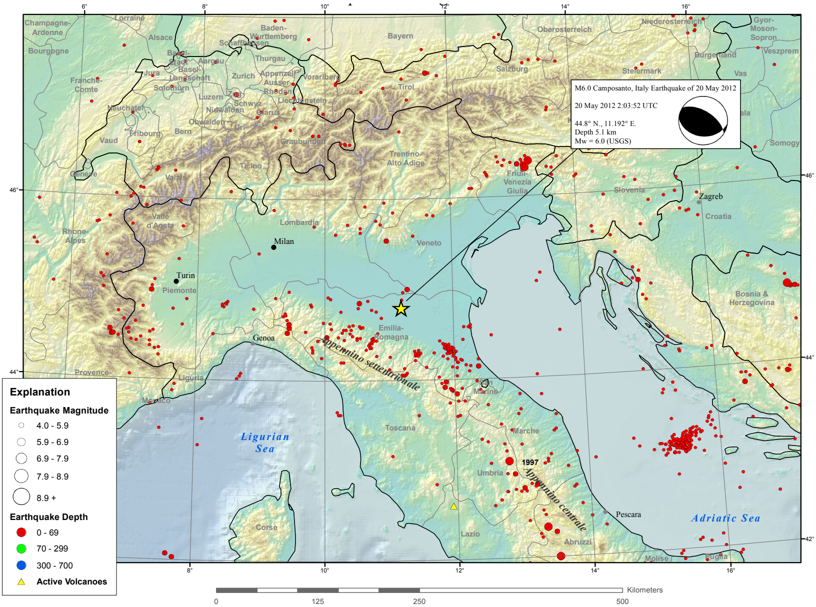


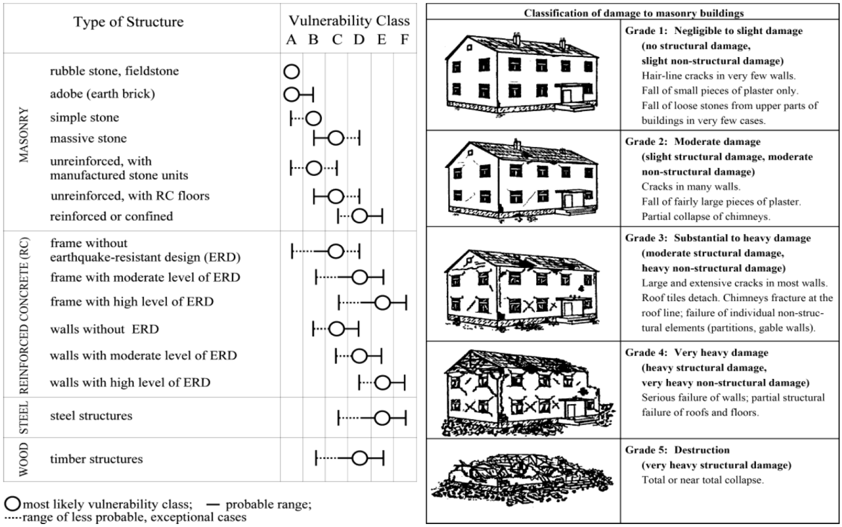
Figure 5.2: Epicentral area affected by the seismic events

Source: USGS Earthquake Hazards Program

instance, a locality is classified with VIII grade when many buildings with high vulnerability level (class A) show very heavy damage (D4) and few of the same buildings (class A) are completely collapsed. Technically, the method is based on the EMS-98 intensity definition: it classifies the buildings into 6 classes of decreasing vulnerability, from A to F (Figure 5.3.2a); then it defines the damage distribution depending on the intensity level. The damage levels are 5 (from D1 to D5) and are based on the structural and non-structural features of the buildings (Figure 5.3.2b). The level D0 means the lack of any damage.

5.3.3 Perception of the risk, reconstruction and funding

The region was not considered a highly exposed seismic zone until 2012. With the exception of the seismic sequence of Ferrara in 1570, Argenta in 1624 and Bologna in 1929 (Vannoli et al., 2015), few other small intensity earthquakes have had an impact on the collective memory of their



(a) (b)

Figure 5.3: Buildings vulnerability and damage level.
Source: Grünthal (1998, pp.14–15)

inhabitants. As a result, the perception of a seismic risk was really small in this area compared with the rest of Italy. In fact, PGA values in this area are, on average, only 20% of those characterizing the nearby Apennine mountain chain. [5] Moreover, the INGV estimated the zone's seismic hazard to be about 0.05 and 0.15 in terms of maximum horizontal ground acceleration rate, up to five times smaller than the one estimated in the Appenini zone in the rest of the Italian peninsula.[6] Accordingly, in this area housing construction was not subject to any specific anti-seismic measure compared with the rest of the country.

In the aftermath of the earthquake the perception of the risk dramatically changed and money has been sent to finance the reconstruction as well as to secure the whole area. The intervention was implemented in two stages. A first phase concerned the urgent operations required to provide first aid, to refurbish buildings and equipments, especially those

[5]See http://zonesismiche.mi.ingv.it/
[6]Source: INGV http://www.mi.ingv.it/pericolosita-sismica.

related to water, electric and drainage system . It was implemented in the very next days that followed the end of the seismic sequence. A second phase aimed at financing a number of projects that were precisely targeted to secure and refurbish buildings in accordance with the new seismic risk. On the whole, 11 million euros have been disposed in the province of Mantua, 122 millions in the provinces of Bologna, Ferrara, Modena, and Reggio Emilia, and 8.8 in the province of Rovigo. Money has been sent not only to finance the reconstruction, but also to secure the whole area and to reinforce the anti-seismic system of the buildings, sometimes also getting the chance to make a better sustainability of the energy consumption so to ameliorate the building system, while increasing the seismic safety and the urban quality.

5.3.4 Government actions after the earthquake

After the seismic events in May and June 2012, the Italian government managed the emergency entrusting this task to the Italian Civil Protection and issued several actions for the reconstruction. The law # 122/2012 contains the main measures enacted to regulate the procedures for the intervention during the emergency and reconstruction phases. First of all, a dedicated "fund for the reconstruction of the earthquake-affected areas" was established. The governors of the affected regions were appointed to coordinate the reconstruction activities of their respective administrative competencies. Moreover, the guidelines for the reconstruction measures were listed, detailed as follows.

i) The legal aims of the reconstruction: the term "reconstruction" does not refer only to the actions devoted to repair and restore the buildings affected by the earthquake, but also to increase the seismic safety and the urban quality.

ii) The objects of the reconstruction: all the private and public buildings, infrastructure and productive building units that suffered damage due to the earthquake in a number of municipalities defined by law as epicentral area.

iii) The public authority responsible for the implementation of the reconstruction operations: the regional governors and the deputy commissioners for the reconstruction, in agreement with other administrative entities such as the provinces and the municipalities.

iv) The eligible interventions that could have been financed: the works necessary to repair the damage caused by the seismic events and to allow the full restoration of the building functionality; the interventions for energy efficiency enabling the reduction of losses and the use of renewable energy sources; the restoring interventions of existing facilities or the complete renovation if the repair is too expensive.

v) The administrative procedure to get the funding.

Among the several laws enacted by the government, we can keep track of the distinction between the funds assigned just to secure safety or reinforce the anti-seismic measures of the buildings and the funds devoted to reconstruction and rebuilding: this specification is available for the schoolhouses as well and this is how we get the information concerning the extra funding the schools received in the aftermath of the earthquake.

5.4 Data sources and summary statistics

For our analysis we consider several sources of data at school, municipal, provincial and regional level. Table 5.1 contains summary statistics of the main variables as described in the remaining part of this section.

5.4.1 Educational achievement and school data

As introduced in Section 5.2, the National Institute for Educational Evaluation of Instruction and Training (INVALSI) is in charge of the Italian students' skills and achievements evaluation[7]. We collect test scores for both Mathematics and Italian language of tenth graders and we track the same schools over six years, from 2010/11 to 2015/16, that is two years before and four years after the earthquake. The test scores range between a minimum of 10 and a maximum of 100 reached when all the answers are correct. As the test scores are provided at student level, we aggregate them so to construct several variables at state high school level, for both

[7] We refer to Angrist et al. (2014) and Battistin and Meroni (2016) for a thorough description of the test and a more comprehensive overview. Battistin and Meroni (2016) also offer a novel study on instruction time and students' performance in Italy, using the same kind of data.

mathematics and Italian test scores. The variable "*Score (mean)*" measures the average result of the tenth graders for each school. Then we measure the average test scores of the low-achieving students by considering the fraction of students in the 5th and 10th percentile of the score distribution, so that we have respectively "*Score (p5)*" and "*Score (p10)*". In a similar way, we measure the average test scores of the high-achieving students by considering the fraction of students in the 90th and 95th percentile of the score distribution, so that we have respectively "*Score (p90)*" and "*Score (p95)*".

We obtain INVALSI data also on the share of male and native students and on the cohort size referring to the tenth grade of each state high school.

5.4.2 Earthquake damage assessment and seismic hazard

We use the INGV macroseismic survey which provides estimations of the volume of buildings with a certain level of damage in a given municipality. It matches information from the macroseismic intensity values and the level of vulnerability of the buildings in the municipality, that varies across six classes of vulnerability in relation to the structural characteristics of the buildings. The INGV macroseismic survey is based on a sub-municipal territorial classification provided by the "Italian Revenue Agency" and kindly shared with us by Meroni et al. (2017): for more details we refer to Galli et al. (2012).

We also collect the data for the variable "*Seismic hazard (PGA)*" from the following INGV official website http://esse1-gis.mi.ingv.it/s1_en.php. The seismic hazard is the earthquake occurrence probability within a given window of time and in a given geographic area. Several aspects are taken into account in the hazard assessment: source and patterns of earthquake occurrence, soil types and groundwater conditions. The hazard estimation leads to a risk assessment that affects the land use planning and the building and infrastructure projects. Specifically, we focus on the horizontal peak ground acceleration (PGA), as it is one of the shaking parameters displayed in the probabilistic seismic hazard maps of the Italian national territory for which disaggregate values are available (municipal level). It is commonly used as an index for seismic hazard intensity, i.e. the higher the PGA the larger will be the intensity of possible earthquake in a specific geographic area: accordingly, the higher the PGA, the higher will be the probability to suffer damage on physical infrastructure and buildings. The unit of measure is the gravity acceleration and it refers to the maximum ground acceleration during the earthquakes. For each PGA evaluation the distribution of the 50th percentile is available, for nine different probabilities of exceedance in 50 years (we consider the 10% probability). In our sample, the PGA varies between 0.087 and 0.207, with an average intensity of 0.155.

5.4.3 School funding data

The presidents of the quake-affected regions were appointed as deputy commissioners to promote interventions to reconstruct and to secure the affected areas. Accordingly, several legislative acts were enacted providing the guidelines of such interventions and, among others, also those ones related to the school funding. As explained in Section 5.3, the funds were intended for reconstruction, securing the whole area and reinforcing the anti-seismic system of the buildings, accompanied by an increase of the seismic safety and urban quality in terms of energy consumption sustainability. Both public and private buildings belonging to the municipalities defined by law as the epicentral area were eligible for extra funding. The Italian government made available more than 24.4 millions of euros to several state buildings in 236 municipalities, including 276 state high schools, with the aim of reconstructing damaged buildings, renewing and maintaining all school buildings safe from future seismic threats. From the legislative acts we get the information on the funding received by the state high schools located in the quake-affected areas. These documents report the name of the school, the imputed total amount of money, a short description of the required intervention, the municipality where they belong to and the body responsible for the implementation of the measures, among others. We collect information on 68 state high schools which reported no damage but they were in municipalities that received about 3.6 million of euros to improve the quality of the school buildings. Summary statistics show that these schools received on average 198 euros per student, about 100% of the annual amount in capita expenditure spent in 2013 in Italy (OECD, 2016).[8] This group of schools has been complemented by 105 additional ones, as belonging to adjacent municipalities that had neither been hit by the earthquake nor received extra-

[8] According to the OECD report, the average total spending per student in Italy in 2013 was 9,174 euros; but only 2% (i.e., 184 euros) was devoted to school capital. This amount is very small if one compares it with that funding transferred to schools in other European countries of the same size: capital expenditure in Germany, for example, was about 1,300 euros, and about 1,200 euros in France.

funding according to the enacted measures. To sum up, our sample has 173 state high school and it covers 43 municipalities: for further explanations on these two groups of schools, we refer to the next section. Additionally, we geolocate each school of our sample in sub-municipal areas through a map navigation as provided by the "Italian Revenue Agency", so to be able to match this information with the macroseismic survey one.

Furthermore, we can keep track of the way funds have been employed in the reconstruction carried out in each school: this information is publicly available upon request to the responsible province and it is useful to detect whether or not the interventions in the schoolhouses were invasive.

Table 5.1: Summary Statistics

	mean	sd	min	max	count
Panel A – Treatment and IV					
Spending dummy	0.39	0.49	0.00	1.00	173
Funds per capita ($\times$ 10)[a]	7.79	20.58	0.00	161.29	173
Funds per capita ($\times$ 10)[b]	19.82	29.07	1.45	161.29	68
Seismic hazard (PGA)	0.16	0.03	0.09	0.21	173
Panel B – Mathematics					
Score (mean)	3.80	0.33	2.79	4.46	692
Score (p5)	3.09	0.62	0.00	4.30	692
Score (p10)	3.28	0.50	0.00	4.32	692
Score (p90)	4.15	0.28	3.11	4.59	692
Score (p95)	4.22	0.26	3.11	4.59	692
Panel C – Italian Language					
Score (mean)	4.10	0.27	1.59	4.50	696
Score (p5)	3.57	0.60	0.00	4.39	696
Score (p10)	3.73	0.47	0.00	4.44	696
Score (p90)	4.34	0.20	1.59	4.58	696
Score (p95)	4.38	0.18	1.59	4.59	696
Panel D – Controls					
% Male	0.56	0.26	0.00	1.00	692
% Native	0.82	0.14	0.21	1.00	692
Cohort Size	88.57	77.39	3.00	372.00	692

Notes: [a] All sample. [b] Only treated.

5.5 Empirical strategy

The aim of this work is to detect the impact of school resources on students' achievement. However, the educational outcome and the school spending levels are potentially simultaneously determined. For this reason, we look for exogenous variation in infrastructure spending and we address the endogeneity issue by using data on school funding provided after a natural disaster, namely the Northern Italy earthquake in 2012. The seismic events represent an external shock that enables us to compare similar schools with different entitlement for extra-funding and they ensure random selection for funding eligibility. Accordingly, we implement two intertwined yet different identification strategies: the first one exploits the information on the allocation process, that is whether schools received funding or not (difference-in-differences method); the second one uses the amount of funding that each school received as function of pre-determined seismic risks (IV approach).

Before going into the technical details of each method as presented in Section 5.5.1 and in Section 5.5.2, we would like to clarify some concerns that might arise regarding the empirical framework depicted so far. Similar reasoning can be found in Cipollone and Rosolia (2007), despite the fact that it applies to a different context and refers to a different empirical analysis.

First of all, the students' performance might have been affected not only by the additional resources the schools received, but also by other quake-related shocks. Given the massive collapse of buildings and devastation in few quake-affected areas, the earthquake could have had an impact on the students not only because of the extra-funds provided in the aftermath of the earthquake, but also because of direct consequences in the students' daily life: this could have been the case if, for example, students were not allowed to enter the school or they could not attend class regularly. Second, as the earthquake affected a wide area, the economic environment where the schools are located might have been different, adding confounding factors in the school spending management and in the impact of additional resources in the school system. Third, the earthquake might have played a direct effect on individual schooling: for example, parents might have been concerned to send their children to school, leading to direct consequences on the students' performance. Moreover, the earthquake might have affected the students' performance even in the area not directly involved in the seismic events, by means of geographic spillovers: for example students might have been reallocated in safer schools so that the learning environment would have experience a dramatical change.

To rule out the concerns presented above, we focus on the least affected areas and on schools whose buildings had no damage at all. We use information on the volume of damaged buildings in each municipality as estimated by the INGV in the aftermath of the seism using a macroseismic survey. We only select municipalities where the level of damage of their buildings was assessed by the INGV as "negligible" (D1) or null (D0): for a thorough description of the macroseismic survey and the levels of damage we refer to Section 5.3. In more details, we collect data for a total of 236 municipalities, as shown on the map in Figure 5.5. Out of 236, 69 are discarded as they had a level of damage greater than D1 (see grey shaded areas in Figure 5.5): we keep only schoolhouses located in municipalities where hair-line cracks in the walls or small pieces of plaster broken off could not possibly affect negatively the learning process of the students. Out of the 167 remaining municipalities, only 43 have at least one high school, for a total of 173 schools (white dots in Figure 5.5). The treated schools are those located in treated municipalities (shaded areas in Figure 5.5) and make up a good portion of the total number of schools, 39% (68). Our control group is made up of 105 schools that received no extra-funding and were not affected by the earthquake, but they are located in municipalities proximate to the treated areas (i.e., they either share borders with the treated areas or there is no more than one municipality between them and the treated areas, see dashed areas in Figure 5.5). The map also contains information on the PGA values. The colour bar shows the gradient of PGA for each municipality, from low to high. The amount of extra funding per student in the treated areas was mostly driven by the necessity of safeguarding school buildings from future seismic threats and minimize potential damages to school infrastructure. [9] Hence it is a function, among other things, of PGA levels.

The geographic proximity across the two groups of schools rules out the possibility of having additional shocks stemming from the economic environment. Moreover, as both the school groups are located in the surrounding earthquake zones, both treated and control schools are in not damaged at all areas: in this way we avoid the possibility that direct effect across the two groups could have been different. It is also worth recalling that the seismic events occurred in May–June, so almost at the

[9] See the first decrees enacted by the deputy commissioner, i.e., ODC #2 (16 June 2012) and the ODC #4 (3 July 2012).

end of the school year: as the next school year started in mid-September, there was enough time to ensure school safety and promote calmness among parents.

As for the presence of possible geographic spillovers, by law (e.g. #ODC July 2012, 25) new schools built from scratch or prefabricated school buildings needed to be ready at the beginning of the new school year in place of the schools heavily damaged or even destroyed by the earthquake. Therefore, kids were not supposed to go to surrounding schools. As an additional evidence to this argument, the trend for the average number of enrolled students (10th graders) shows that there is no significant different behaviour after the earthquake events between the two groups (Figure 5.4).

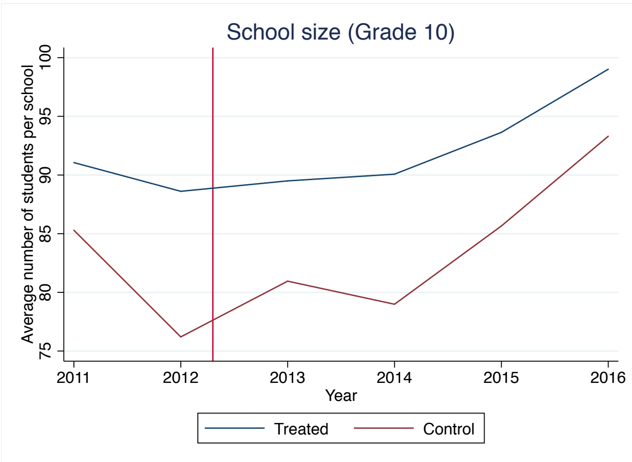


Figure 5.4: Average number of enrolled students (treated and control group)

Finally, Tables 5.2 and 5.3 show the pre-treatment differences in test scores and covariates, respectively, between treated and control group. As we can see, there are no remarkable differences between the treated and the control schools for the mathematics test scores and for the students' characteristics across the two groups.

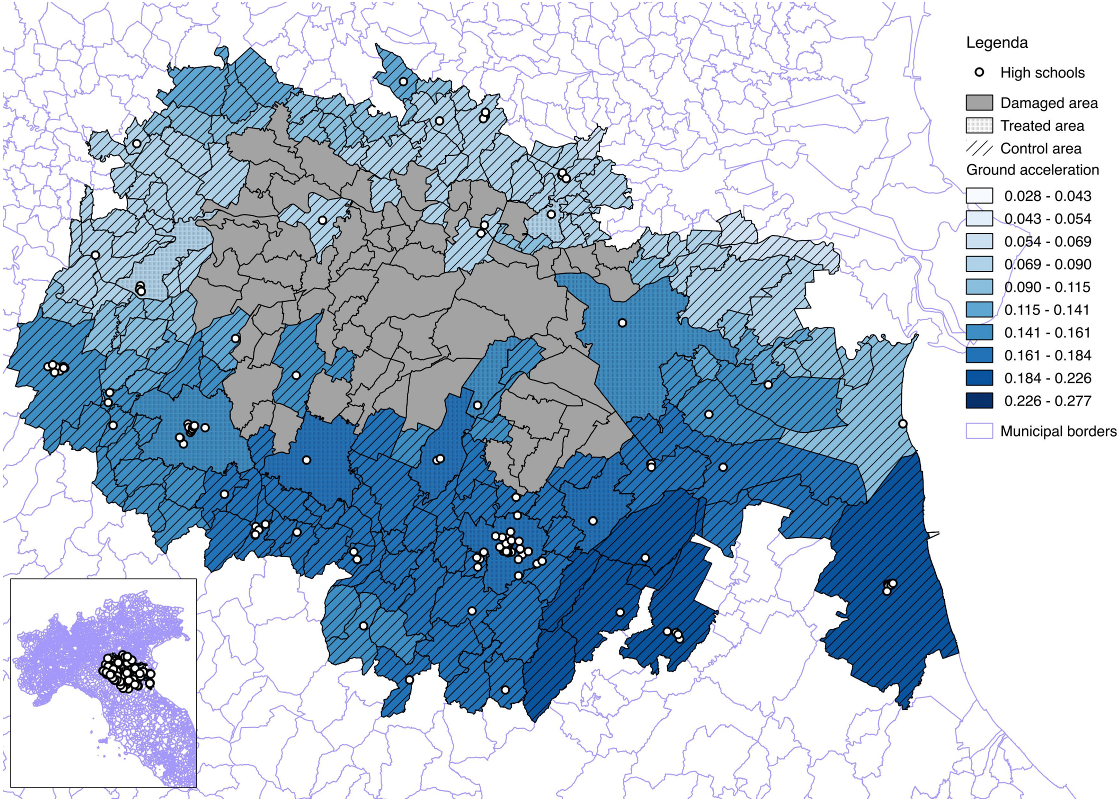


Figure 5.5: Treated and control areas

Table 5.2: Pre-treatment Test Scores

	Score (mean)	Score (p5)	Score (p10)	Score (p90)	Score (p95)
Panel A Mathematics.					
T-C	1.575	0.803	0.620	2.370	1.908
	(2.177)	(1.840)	(1.925)	(2.535)	(2.621)
Control	47.490***	27.683***	32.177***	63.329***	67.599***
	(1.376)	(1.164)	(1.195)	(1.618)	(1.652)
Panel B Italian language.					
T-C	3.955**	5.298**	4.303*	3.023**	2.322*
	(1.899)	(2.612)	(2.464)	(1.456)	(1.349)
Control	66.643***	45.274***	51.292***	80.476***	83.372***
	(1.295)	(1.636)	(1.585)	(1.078)	(0.989)
Observations	270	270	270	270	270

Standard errors clustered at the school level in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5.3: Pre-treatment Covariates

	% Males	% Natives	Cohort size
T-C	-0.067	0.017	7.961
	(0.044)	(0.025)	(13.058)
Control	0.589***	0.826***	80.848***
	(0.027)	(0.015)	(7.332)
Observations	270	270	270

Standard errors clustered at the school level in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.5.1 Difference-in-differences estimation strategy

Our identification strategy is twofold. First, we use the quasi-experimental setting induced by the 2012 Northern Italy earthquake to get a handle on the direction of causation in the infrastructure spending – students' achievement relationship: in other words, we investigate the treatment effect of receiving or not additional funding. Using information from the map in Figure 5.5, we can measure the impact of receiving additional resources on test scores by comparing the evolution of test scores before and after the allocation of funds in the recipient areas as compared to those that did not receive extra-funds. We start with a simple empirical research design, a difference-in-differences estimation strategy, which takes the following form:

$$\begin{aligned} log\, y_{it} = \alpha_0 + \alpha_1 D_i + \alpha_2 P_{t-1} + \alpha_3 D_i * P_{t-1} + X_{it}'\alpha_4 + \\ + \mu_i + \eta_p * P_t + \theta Trend + \varepsilon_{it} \end{aligned} \tag{5.1}$$

where the outcome variable y_{it} denotes the average test score in either mathematics or Italian language in school i in year t; D_i is a dummy that takes value one if the school belongs to the treated area; P_t is a dummy that takes value one if the observation is in the post-treatment period (i.e., post 2012);[10] X_{it} is a vector of school covariates which includes the school size, the shares of male as well as the share of native students in each school; μ_i is the school fixed effect, which absorbs school-specific constant (or slow-moving) features; as provinces could have implemented local interventions after the earthquake, we interact province fixed effect η_p with P_t to control for province-specific policies after 2012;[11] θ is the coefficient of a school-specific time trend variable and ε_{it} is an error or disturbance term. $D_i * P_t$ is the interaction between the treatment schools D_i and P_t, the dummy variable equal to one in the post-treatment period; therefore, α_3 is our parameter of interest, the difference-in-differences estimates of the impact of receiving funding on students' achievement. Note that, for small values of the coefficient, 100*α_3 can be interpreted as the percentage increase in the test score when schools receive extra funding.

[10] We lag the treatment by one year to allow time for the funding to be invested.

[11] A province is an administrative division between a municipality and a region, and constitute the third NUTS administrative level. Provinces have, among other functions, the local planning and the coordination of schools activities. In our sample, we have a total of 10 provinces.

5.5.2 Instrumental variable strategy

Second, we want to offer estimates of the elasticity of test scores with respect to spending per capita: in other words we want to explore the treatment intensity effect depending on the amount of received funding. Yet, as noted above, idiosyncratic changes in school spending are likely endogenous as the amount of funding allocated to each school can be correlated with unobservable school-level characteristics. To quantify this relation, we estimate 2SLS models where we instrument for school spending with the values of peak ground acceleration (PGA), the maximum ground acceleration during the earthquakes. Recall that funding was allocated to schools to reduce the vulnerability of their buildings to earthquakes and more funding per capita was granted to schools in municipalities with higher earthquake risks. The proposed instrument is thus strongly correlated with school funding. At the same time, it is uncorrelated with school-level unobservables that might affect test scores. Thus, PGA offers a valid instrument.

The second stage of the IV estimation is given by:

$$log\, y_{it} = \beta_0 + \beta_1 \widehat{FUND}_{it-1} + X_{it}'\beta_2 + \mu_i + \eta_p * P_t + \theta Trend + \varepsilon_{it} \quad (5.2)$$

where the outcome variable y_{it}, the vector of controls at the school level, the trend variables and the fixed effects are the same as in equation (1). $\widehat{FUND}_{it}$ is the estimated funding per pupil as predicted by the first stage. The equation we estimate in the first stage uses the PGA level in the area where the school is located as an instrument for actual funding. Given the log-linearity of the model, the interpretation of β_1 is that of a proportional change in the test score given a unit change in funding, holding all else constant.

5.6 Results

In Table 5.4 we present the relation between funding and student scores in mathematics, whereas in Table 5.5 we focus on Italian language. In column 1 of each table we use as dependent variable the average score for all students, in columns 2 and 3 the test scores for students in the 5th and 10th percentile of the score distribution (i.e., low-achieving students), and in columns 4 and 5 the test scores for students in the 90th and 95th percentile (i.e. high-achieving students).

In panel A we show a naive OLS estimation, which reveals a positive correlation between funding per pupil and test scores. If for purely illustrative purposes one interprets the OLS estimates as causal, then, according to the estimates, a one-unit increase in school infrastructure spending per student (that is, 10 euros) is associated with an estimated increase in test scores in mathematics in the range of 0.1% to 0.7%, holding all else constant. The relation is insignificant at conventional levels when we replace test scores in mathematics with those in Italian language (Panel A, Table 5.5).[12]

In panel B we turn to our quasi-experimental design and we uncover a positive effect of receiving extra funding on test scores, although the relation is still not significantly different from zero for Italian language. In more detail, test scores increase by 10% overall if a school is a recipient of funding, and the effect is substantially larger for low-achieving students (between 26% and 33%).

Turning to the elasticity of student outcomes with respect to the amount of resources devoted to school infrastructure, recall that in panel A our main coefficients of interest are most certainly contaminated by endogeneity from uncontrolled confounding variables. Therefore in panel C we turn to the estimated coefficient of school funding in the second stage of our 2SLS. We use the PGA, an index of seismic hazard, as exogenous instrument. As we can see, the coefficients are now substantially larger than those of the naive regressions in panel A and they are all statistically different from zero. Distributing an extra 10 euros per pupil to schools will produce an estimated test score gains in mathematics in the range of 0.7% to almost 6.3%.[13] Again, we find that the marginal return to investment in school infrastructure is greater the lower the grade of the students. Interestingly, we now obtain similar results with test scores in Italian language and the estimated magnitudes of the relationship between funding and students' achievement are not only statistically significant but also economically meaningful.

In panel D we show the reduced form and the first stage estimates. As expected, we find that an increase in the PGA level has a sizable im-

[12]Note that all models include the share of males, of native students and the total number of students in each school as well as school fixed effects, time trends and interactions between province fixed effects and post-treatment period dummy. Using linear trends, quadratic trends, cubic polynomial in time (i.e., t, t^2, and t^3) or year dummies produce similar results.

[13]These results are not driven by the upper tail of funds and are robust to the exclusion of the top 5% of the schools from the sample, i.e. those that received more than 800 euros per student.

pact on students' scores. At the same time, the first stage reveals that the PGA level leads to a higher amount of infrastructure funding received by the school. We report the Kleinbergen-Paap F-Statistic, which is similar to the conventional F-statistic, but takes into account the clustering of the standard errors. The values are all above conventional levels characterizing weak instruments.

To dig deeper into the relationship between school funding and students' standardized test scores, Figure 5.6 shows the relation between the estimated coefficient β_1 in equation 5.2 and the quantiles of the distribution of the test scores. As the figure clearly reveals, allocating additional funding to schools' infrastructure has higher marginal effects on the achievement of students with the lowest scores on the standardized tests. Whereas in Italian language the pattern is less clear-cut, in mathematics the estimated effect decreases monotonically as we move from the 10th to the 90th percentile of the standardized test score distribution. Results are overall similar when we look at relation between the estimated coefficient α_3 in equation 5.1 and the quantiles of the distribution of test scores (see Figure 5.7). We can conclude from these two tables that the previous results using a difference-in-differences approach are strongly borne out by this new set of empirical results. The effect of school funding on students' achievement is overall quantitative large, statistically significant and robust, in particular in mathematics and for low-achieving students.

Table 5.4: Secondary School, Mathematics: funding and students scores

	(1)	(2)	(3)	(4)	(5)	(6)
	Score (mean)	Score (p5)	Score (p10)	Score (p90)	Score (p95)	Funding p.c. (log)
Panel A OLS Estimation.						
Funding p.c. (log)	0.002***	0.007**	0.005**	0.001**	0.001***	
	(0.001)	(0.003)	(0.002)	(0.000)	(0.000)	
Panel B OLS Estimation.						
Spending dummy	0.099***	0.329***	0.262***	0.043**	0.040*	
	(0.024)	(0.094)	(0.063)	(0.021)	(0.021)	
Panel C IV Estimation – Second Stage.						
Funding p.c. (log)	0.023***	0.063***	0.054***	0.009***	0.007***	
	(0.005)	(0.014)	(0.012)	(0.003)	(0.002)	
KP F-Statistic	17.324	17.324	17.324	17.324	17.324	
Panel D Reduced Form and First Stage.						
Seismic hazard	1.204***	3.284***	2.803***	0.495***	0.381***	52.255***
	(0.101)	(0.441)	(0.291)	(0.093)	(0.099)	(12.140)
Observations	692	692	692	692	692	692
N_i	173	173	173	173	173	173

Notes: School Fixed-effect models. All regressions include fraction of males, fraction of native students, and number of students in the tenth cohort as well as linear trend and province dummies interacted with P_t. Funding per student are expressed in 10 euros. Standard errors are clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5.5: Secondary School, Italian: funding and students scores

	(1)	(2)	(3)	(4)	(5)	(6)
	Score (mean)	Score (p5)	Score (p10)	Score (p90)	Score (p95)	Funding p.c. (log)
Panel A OLS Estimation.						
Funding p.c. (log)	0.000	0.003	0.001	0.000	0.000	
	(0.001)	(0.003)	(0.002)	(0.001)	(0.000)	
Panel B OLS Estimation.						
Spending dummy	0.020	0.073	0.051	0.008	0.005	
	(0.022)	(0.064)	(0.042)	(0.019)	(0.019)	
Panel C IV Estimation – Second Stage.						
Funding p.c. (log)	0.006**	0.011*	0.008**	0.004*	0.005**	
	(0.002)	(0.006)	(0.004)	(0.002)	(0.002)	
KP F-Statistic	14.771	14.771	14.771	14.771	14.771	
Panel D Reduced Form and First Stage.						
Seismic hazard	0.328***	0.623*	0.462**	0.218**	0.267**	54.827***
	(0.113)	(0.345)	(0.197)	(0.106)	(0.109)	(13.773)
Observations	696	696	696	696	696	696
N_i	173	173	173	173	173	173

Notes: School Fixed-effect models. All regressions include fraction of males, fraction of native students, and number of students in the tenth cohort as well as linear trend and province dummies interacted with P_t. Funding per student are expressed in 10 euros. Standard errors are clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

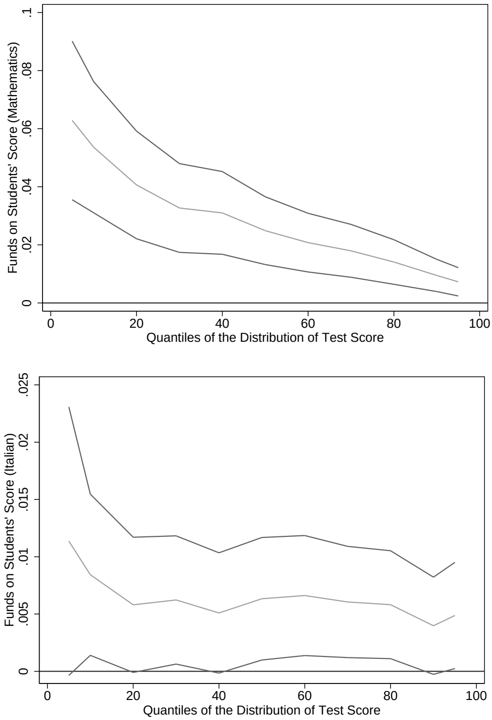


Figure 5.6: Estimated impact of school funding on test scores by quantiles of the distribution of test scores

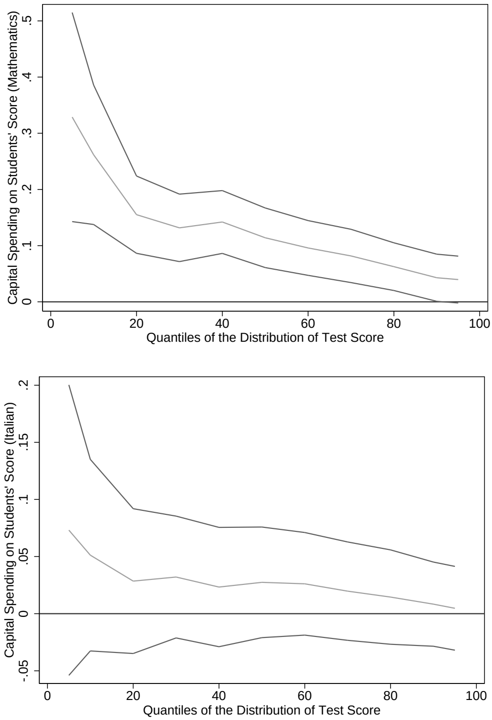


Figure 5.7: Estimated impact of receiving extra-funding on test scores by quantiles of the distribution of test scores

5.7 Conclusion

In this chapter we explore the impact of school infrastructure investments on students' achievement. We use data on school funding provided after a natural disaster, a magnitude 5.9 earthquake that hit the Northern part of Emilia Romagna region in May 2012, affecting an area of 3,500 squared kilometers. We use information on the allocation process (whether schools received funding or not) and on the amount of funding that each school received (function of pre-determined seismic risks) to implement two intertwined yet different identification strategies, so as to give our regression estimates a causal interpretation.

Our empirical results suggest that doubling school infrastructure spending reflects onto an increase of students' test score up to 6.3% for mathematics and low-achieving students. A set of facts, peculiar to the Italian school system, may help us reconciling our findings with recent contributions that specifically use US data. Contrary to the United States, few resources are spent in school capital in Italy, about 184 euros per student in 2013, which places Italy near the bottom of school infrastructure spending (OECD, 2016). Whereas the average condition of school infrastructure is quite poor (by one estimate, more than 39% of school buildings need urgent maintenance, see e.g., Antonini et al., 2015) interventions on school facilities are likely to affect the health, safety and morale of students and teachers and in turn their ability to learn and teach. As such, our study outlines the role of physical capital spending in improving the learning environment of high schools and offers potential policy prescriptions for investing in school infrastructure.

Chapter 6

Conclusions and Discussion

In line with the subsidiarity principle, services should be provided by the government closest to the citizens. Local authorities have the potential to better recognize and respond to local needs. They can provide services closer to the preferences and tastes of their citizens and more quickly adapt to changing economic conditions (Slack and Bird, 2013). In this way, services are delivered in a more targeted way and without an imposition coming from higher levels of government, pursuing a greater degree of efficiency in the management of increasingly limited and scarce public resources (Da Cruz and Marques, 2014). Moreover, closer interaction between government and locals enables higher citizen involvement in the local decision-making process, ensuring accordingly a clearer accountability and fostering a better combination of taxes and services in terms of local competition (Asatryan and De Witte, 2015).

The concept behind this principle has been acknowledged worldwide, at sub-national, national and international level. At European level, the subsidiarity principle is one pillar of the functioning of the European Union, since the Treaty of Maastricht signed in 1992. Article 5 of the 'Treaty on European Union' clearly states this principle, even if the idea was already present in the 'European Charter of Local Self-Government' adopted by the Council of Europe few years before, in 1985. In the Charter, "*local self-government denotes the right and the ability of local authorities, within the limits of the law, to regulate and manage a substantial* *share of public affairs under their own responsibility and in the interests of the local population*" (Article 3.1) and "*public responsibilities shall generally be exercised, in preference, by those authorities which are closest to the citizen*" (Article 4.3), underlying the importance of the local form of government. The crucial role of local responsibilities has been acknowledged even at a more supranational level. In fact, local governments have been identified as the key to meeting the Sustainable Development Goals (SDGs) adopted by the UN General Assembly (2015) and to be achieved by 2030. These goals deal not only with vital challenges, such eradicating poverty, fighting gender inequality and tackling climate change, but also with public goods provision like education, water services and health, among others. Despite addressing global issues, they are directly relevant to citizens' daily lives and for this reason more tailored solutions should originate from local action. Accordingly, sub-national governments are considered as fundamental in fostering citizen participation and in bringing higher forms of government closer to the citizens for SDGs' achievement. Across countries, there is a huge degree of heterogeneity in the size and number of sub-national governments (see for example OECD 2017d). Accordingly, while monitoring the services and assessing the efficiency and effectiveness of the public expenditure at local level, the considered evaluation tools should be flexible enough to tackle different priorities and should take into account the heterogeneity within and across country.

In this dissertation, modelling tools and empirical applications for local level analysis are provided. In particular, the analysis of two complementary aspects are taken into account, namely service provision and resource management. On the one hand, as outlined above, local governments are in charge of providing many services, covering a broad variety of intervention areas, such as educational and social care services, water sanitation and waste disposal management, local security and housing services, among others. Therefore, tools are required to measure the overall level of provided services encompassing different functions or alternatively to focus on a specific sector further investigating related issues, for example questioning whether its provision is environmentally sustainable or not. On the other hand, local governments face budget constraints and limited resources. Therefore, the way resources are spent for delivering local services cannot be ignored and its analysis should go in pair with the service provision assessment, checking whether public expenditure is managed in compliance with the principles of efficiency and effectiveness. Chapter 2 and 3 deals with the service provision analysis, encompassing first all the local tasks and then focusing on a specific area, water management. Water is still not accessible to everyone and its management has some peculiar characteristics which make it worth to be explored: the presence of huge infrastructural costs has lead to an increasingly significant public-private partnership, which in turn has raised the attention on performance evaluation not only in terms of economic and financial profitability, but also in line with the environmental sustainability aspects for the sake of the people and of the planet. In a symmetric way, Chapter 4 and 5 address the public spending analysis, providing first an overall assessment across the main local competencies and then focusing on a specific area, the education sector. Despite funding in the education sector representing a significant expenditure item on the overall level of public spending, its effect on the students' achievement is still debated: in a context where the lack of investment in school infrastructure is an acknowledged issue worldwide and calls for more school capital funding, this topic turns out to be relevant and worth to be further investigated.

The measurement of municipal service provision is a complex task. Local government competencies cover several intervention areas and there is a wide degree of heterogeneity among municipalities in their political preferences and in their local characteristics. To encompass all these aspects in a single index, **a dynamic robust conditional directional distance function Benefit-of-the-Doubt Composite Indicator with ARI restrictions** is proposed in Chapter 2. Overall, the approach ensures an objective way to determine how each municipal area enters in the evaluation, while granting the most favourable aggregating scheme for the units under analysis. The information on the municipal expenditure composition is included through the weight restrictions specification. The directional distance function formulation makes possible evaluation even along undesirable features, which should be reduced rather than maximized in the problem. The robust conditional version of the model controls for the municipal operating context and the time dimension. The proposed composite indicator is suitable not only for benchmarking and detecting best practises (OECD, 2008), but also for further investigating other aspects. First of all, we can explore how municipal characteristics influence overall service provision through statistical inference, detecting whether the background condition inclusion favours or not the assessment. More broadly, the obtained composite indicator can be used to explore the relationship between the provided municipal services and some relevant issues in the municipal management, as for example government size expressed in terms of the tax burden imposed on citizens, which pay the taxes for the local public goods they receive.

When evaluating the public good supply, the quality of the provided services and their environmental impact should be considered as well. These issues are relevant especially in particular areas, as for example services related to the water industry. Accordingly, in Chapter 3 **an integrated approach that combines insights from the Non-radial Directional Distance Function (NDDF) and the Analytic Hierarchy Process (AHP) in a Water Performance Index** is proposed. As from the first technique, it is possible to simultaneously reduce inputs and bad outputs while expanding good outputs to different extents. On the other hand, in line with the AHP approach, the preferences of the decision makers (e.g. water utility managers and water authorities) can be directly included in the weights' specification, determining the relative importance of inputs, good and bad outputs. Specifically, a particular emphasis can be assigned to those units more "environmentally" focused and that would have been otherwise penalized for the higher costs they face for keeping their water quality commitment. The obtained water performance index is suitable for further exploring the impact of external factor characteristics, so to give the water authorities and the decision makers additional insights for the water management service under analysis.

In public sector analysis, attention is not only on the service provision side, but also on the way resources are spent in terms of efficiency and effectiveness. To evaluate local government expenditure efficiency considering municipal competencies both one by one and all together, **a three-stage Data Envelopment Analysis (DEA) model** is introduced in Chapter 4. The proposed model is flexible to focus first on each municipal area and then to perform an overall analysis aggregating in a composite indicator the scores obtained in the previous step, according to two different ways. The first one follows the DEA approach using a common system of weights to grant the comparison on a common ground; the second one relies on the municipal expenditure composition information available in the balance sheets. Despite being different, these two indexes are complementary tools to provide useful insights to policy-makers, enriched by further investigation of municipal features' impact on the estimated level of efficiency.

As the goal of policy-makers is not only "doing things right", but also "doing right things", resource management needs to be evaluated also in terms of effectiveness. In particular, spending in the education sector deserves to be carefully considered as it represents one of the most important expenditure item in public spending. More specifically, in Chapter 5 the **endogeneity arising while assessing school resource effectiveness is tackled by two complementary quasi-experimental designs, the Difference-in-Differences and the Instrumental Variable strategies**. The advantage of this intertwined approach is twofold: first, it is apt to get whether receiving funds affects educational achievement; second, it measures the intensity effect of receiving capital investment on student outcomes.

6.1 Evidence-based policy implications

To address specific questions relevant from a policy point of view while showing the potentiality of the novel techniques proposed in the thesis, we consider study cases from Belgium and Italy.

Specifically, in Chapter 2 the dynamic robust conditional directional distance function Benefit-of-the-Doubt Composite Indicator with ARI restriction is estimated for 307 Flemish municipalities over 2006–2011. The composite indicator scores show the methodological importance of the proposed integrated and fully flexible analysis, as it creates a level playing field among the municipalities under analysis. First of all, when considering the operating environment each municipality has to work in and the priorities to be aligned with, it turns out that there is only little room for municipal service provision improvement. Therefore, when benchmarking and detecting the best practices, the municipal characteristics should be strongly considered to grant a fair comparison. Moreover, the background variables can be further explored in the way they affect overall service provision through statistical inference, detecting whether their inclusion favours or not the assessment. For example, in line with the literature (see for example Kornai et al., 2003; Šťastná and Gregor, 2015), the level of municipal fiscal income plays an unfavourable role in the service provision, pointing at the so-called "wealth effect", so that the higher the level of available resources, the higher their potential mismanagement. This suggests a more stringent control over those richer local governments that have the potential to provide even more services to the citizens (or of better quality). As concerns the share of elderly people and foreigners, a favourable influence is found, possibly suggesting that the provision of the related services to a greater catchment area would benefit from scale economies. Looking at the political component, the findings confirm the idea that a more left-wing government favours a higher overall level of municipal activities. Beyond these considerations, the relationship between the composite indicator scores and the tax burden is explored to test the existence of an optimal government size, for the first time at municipal level. In this context, the tax burden is measured as tax revenue over local taxable income. This topic is still highly debated and it affects not only policy decision making but also the citizens, directly involved as they pay the taxes for the local services they receive. In the presented application, an “inverted U-shape” is recurrent and significant in every model and weight restriction specification, confirming the hypothesis proposed in the economic literature: beyond an optimal point, despite a higher level of revenues, a lower level of public goods is provided to the citizens (De Witte and Moesen, 2010). The estimated optimal government size is the same as the current level of local tax rate if considering the very basic and unrestricted model. However, when including the economic variables and the expenditure structure, the average optimal size estimated for Flemish municipalities increases at 5.29%, 1,54% higher than the current tax rate in Flanders. This shows room for increasing local taxation of those local governments adopting a lower level of tax rate so to be able to afford a higher level of service provision, observing that municipalities with similar characteristics can do it. As an alternative to the tax rate increase, this evidence might point to a further need of municipal aggregation so to make wider the tax base and to share the costs of the services among a higher number of taxpayers (Slack and Bird, 2013).

In Chapter 3 the integrated Non-radial Directional Distance Function/Analytic Hierarchy Process model is estimated to compute Water Performance Indexes for 96 Italian wastewater treatment plants (WWTPs) in 2014, including nitrogen as one of the most harmful pollutants to be removed in the outgoing treated water. The environmental efficiency scores are explained by means of several variables related to the technical features of the WWTP. The obtained results could provide useful suggestions for the water utilities, environmental agencies and regulators in terms of policy implications (Akhmouch and Correia, 2016; OECD, 2017a). Generally, the evidence shows the added value arising from an integrated performance assessment, that penalizes the WWTPs aiming only at getting cost savings and achieving poor environmental standards (Molinos-Senante et al., 2014a). Specifically, whenever the expansion of outputs and the contraction of undesirable output are allowed, the efficiency scores are affected by the plants’ ability to respect the legal nitrogen concentration threshold and by the percentage of the discharged industrial wastewater. From the environmental agency side, the introduced performance indexes suggest to conduct inspection activity on small plants treating only domestic sewage. The result is in conflict with prior literature that shows a poor performance among the plants treating the sewage from factories and farms (Guerrini et al., 2016). The novelty of the result obtained in this application can be attributed to the measurement of nitrogen as undesirable output introduced for the first time in this literature and commonly acknowledged as one of the most harmful pollutants in the water (Metcalf & Eddy et al., 2003). The environmental controls concerning the nitrogen concentration regulatory limit should also be increased, so to stimulate a better functioning of those plants that do not respect the set threshold and turn out to work in a less efficient way. Moreover, in line with the literature (e.g. Hernández-Sancho et al., 2011b; Molinos-Senante et al., 2014a,b) the population equivalent size and the estimated dry weather flow have a significant impact on the WWTP performance, irrespective of the model specification. This evidence gives clear indication for water utilities in term of WWTPs' size. The managers should plan to exploit larger scale economies: this would imply higher cost savings, but at the same time higher environmental standards achievement. These findings are in line with the Principles enacted by the OECD on Water Governance, promoting among others the *water management at the appropriate scale(s)* and *co-ordination between the different scales* (Principle 2, OECD, 2015), the *enforcement of rules, procedures, incentives and tools (including rewards and penalties) to promote compliance and achieve regulatory objectives in a cost-effective way* (Principle 7.e, OECD, 2015), the *evidence-based assessment of the distributional consequences of water-related policies on citizens, water users and places to guide decision-making* (Principle 11.d, OECD, 2015).

As for the municipal spending efficiency analysis, in Chapter 4 a three-stage Data Envelopment Analysis model is applied to measure the non-aggregate and the overall local government expenditure efficiency for 282 Tuscan municipalities in 2011. Although the two proposed composite indicators are derived following two very different approaches, the conclusions are basically the same. Several municipal features can be explored to detect their correlation with the level of the estimated efficiency scores: in particular, a way to cluster the municipalities into four groups and to display them in four quadrants is presented, so to investigate which are the municipal characteristics that drive the most efficient performance and to detect possible room for improvement in the efficiency level just intervening in the composition of the expenditure looking at the average. The results are consistent. In particular, some new evidence about the long debated issue of the municipal size comes out. In fact, according to the overall analysis, municipal size really affects the efficiency of public expenditure: the bigger is the municipal catchment area, the lower the average cost in the provision of municipal services, which in turn makes possible to provide more differentiated and complex services (Balaguer-Coll et al., 2007; Doumpos and Cohen, 2014). In particular, inefficiency can be related to the presence of too many small fragmented municipalities: this might suggest an aggregation in terms of either mergers or joint-management of the main functions among the smallest municipalities to exploit scale economies, especially for more capital-intensive services, such as water, sewers, and transportation (Geys et al., 2008; Slack and Bird, 2013). This is in line with legislative measures enacted by the Tuscany region proposing mergers or joint-management among municipalities and, more in general, it represents one of the solutions adopted across the European countries facing the same issue.

To evaluate school resource effectiveness, in Chapter 5 a specific quasi-experimental setting induced by 2012 Northern Italy earthquake is exploited to apply both Difference-in-Differences and Instrumental Variable strategies: 173 high schools from the school year 2010/11 to 2015/16 are considered, sharing similar characteristics but different entitlement to receive extra funds in the aftermath of the seismic events. As for the treatment effect, a positive effect of receiving extra funding on mathematics test scores is found and it is substantially larger for low-achieving students. As for the treatment intensity effect, the evidence shows that distributing an extra 10 euros per pupil to schools will produce an estimated test score gain in mathematics up to 6.3%, specifically for low-achieving students. The findings are a bit in contrast with evidence obtained by US analysis, where little or no effects are found in terms of resource effectiveness on educational outcome (Cellini et al., 2010; Martorell et al., 2016), but this might be related with the school capital spending choice across the countries. As in Italy several investments have been planned to be implemented soon, the findings point to a potential radical change in the trend highlighted by the OECD Pisa report (OECD, 2016b). In fact, low-performance students turn out to perform worse than the average: school infrastructure investment might be able therefore to target in particular this group of student and to have a positive impact on their educational achievement.

6.2 Study limitations and future lines of research

The analysis performed in this thesis offers further insights from both an empirical and a methodological point of view. Some of them are driven by the limitations of the analysis itself.

First of all, the proposed modelling tools are applied to specific contexts, accordingly the general validity of the findings should be considered carefully. Belgium and Italy are the study cases of the present empirical analysis, interesting for their common and specific features, apt to provide complementary insights and informative evidences. Obviously, they have some characteristics that would not be shared by other countries, despite the fact that the topics addressed in this dissertation are commonly questioned at European and even broader international level. This is the reason why, for example, the findings in Chapter 5 are at odds with the evidence obtained by US analysis, where little or no effects are found in terms of resource effectiveness on educational outcome (Cellini et al., 2010; Martorell et al., 2016). Actually, the peculiarity of the Italian school framework does make the obtained insights meaningful and relevant. In fact, differently from the 8% of capital expenditure share spent in the United States, only the 2% is spent in upper secondary schools in Italy (OECD, 2017b), positioning this country close to the bottom of the OECD school infrastructure spending. This different starting point might explain why an influx of cash turns out to be effective in the Italian school infrastructure. However, as the lack of investment in school infrastructure is an acknowledged issue spread across OECD countries (e.g. Australia, Italy, UK, US, Germany), an interesting line of further research might point to exploit the information available from the OECD PISA test scores to detect school resource effectiveness in a broader context. Obviously, the main challenge is to find an insightful quasi-experiment setting to be employed using policy evaluation techniques apt to unravel the endogeneity issue and to provide more general findings.

Furthermore, the obtained results should be interpreted with caution for other two main reasons: they might depend on the model specification and/or on the data choice (Cordero et al., 2016). As concerns the first point for example, the evidence obtained for Chapter 2 and 4 relies on a specific linear aggregating scheme for the composite indicator (CI) construction, that is the Benefit-of-the-Doubt (BOD) approach (Cherchye et al., 2007; Melyn and Moesen, 1991). However, alternative aggregating schemes are available in the CI literature and specifically in the BOD-setting, as for example the *multiplicative aggregation* to name one (for an extensive review, see Rogge, 2018; Van Puyenbroeck and Rogge, 2017; Verbunt and Rogge, 2018). In particular, recent studies tackle the observation-specific optimal weights obtained in the basic BOD model by proposing alternative ways to get a common set of weights and accordingly to make more intuitive the comparison of the units under analysis on a common ground. In Chapter 4, the approach proposed by Bernini et al. (2013) has been considered, but other options might be explored, as for example the one introduced by Tofallis (2013) and more recently by Van Puyenbroeck and Rogge (2017) and by Verbunt and Rogge (2018), to provide complementary insights and check the consistency of the proposed tools and results.

On the other hand, the choice of the variables and, more broadly speaking, the data availability do play a role in the empirical application. For example, in Chapter 3 the wastewater treatment plants' efficiency is addressed together with the environmental sustainability issue, specifically referring to the quantity of nitrogen left in the outgoing water. However, the analysis might take into account also other relevant residuals such as phosphorus, pharmaceutical pollutants, toxic metals and therefore further undesirable outputs might be chosen to enrich the overall evaluation. In the analysis outlined as such, the suggested evaluating tool might be very promising for further environmental efficiency analysis of the wastewater treatment plants and, more generally, of the supplied water services. As concerns the water utility services, water losses might be included as undesirable output, encompassing another issue acknowledged worldwide and directly linked to the ongoing water crisis. Environmental sustainability is not the only key element to be considered in water service evaluation: also customer satisfaction and the quality of provided services are increasingly important (see Romano et al., 2017). Accordingly, data related to this aspect should be included as well so to provide a broader and more insightful assessment of the services under analysis.

More broadly speaking, the relevance of the quality aspect and of citizens' satisfaction is remarkably pervading all public sector intervention areas. This is witnessed for example by the increasing development of on-line platforms to keep track of citizens' needs and complains at municipal level or even to encourage active participation and transparent accountability. The importance of the co-production role of the citizens is more and more emphasized (De Witte and Geys, 2013; Parks et al., 1981; Whitaker, 1980) and several initiatives point toward this direction, as for example the project "CitizenPoweredCities: Co-producing better public services with citizens" at OECD level. Getting access to information related to these topics and improving the proposed modelling tools to include them in the analysis would add a very interesting dimension to the local government evaluation as introduced in Chapter 2 and 4.

The possibility of including more data in public sector analysis points to another important issue concerning both the efficiency and the effectiveness analysis in general: the problem of omitted information and specification bias. Referring to the effectiveness analysis, this problem has been addressed by applying policy evaluation techniques as in Chapter 5. However, as concerns the efficiency analysis, this is still an open issue that cannot be neglected (Cordero et al., 2016; De Witte and López-Torres, 2017). In fact, the presence of potential endogeneity might first cause biased results and second prevent the interpretation of the results in a causal way. This topic started receiving increasing attention in the recent literature (see for example Cazals et al., 2016; Cordero et al., 2015; Santín and Sicilia, 2017; Simar et al., 2016) and it represents a promising line of research. Specifically, an interesting way to tackle this issue is to integrate performance efficiency tools with policy evaluation techniques, combining the insights of the efficiency and the effectiveness analysis. With reference to the Italian context for example, the spending efficiency analysis as introduced in Chapter 4 might be extended accordingly. The causal impact of mergers and/or joint management of municipal functions on the overall spending efficiency level should be investigated by exploiting quasi-experimental setting, controlling before and after the implementation and monitoring those who complied or not with the legislative measures.

Appendix A

Supplementary material for Chapter 2

A.1 Data sources and description

Data are available via `http://statistieken.vlaanderen.be/QvAJAXZfc/notoolbar.htm?document=SVR%2FSVR-alle-domeinen.qvw&host=QVS%40cwv100154&anonymous=true`.

In Table A.1, the list of the variables as downloaded is reported with their original Dutch names. In addition to the listed variables, the data linked to *Population Growth* and *Ideological Complexion of the local Government (ICG)* have been provided by De Witte and Geys (2009).

Per capita variables as listed in Table 2.1 have been obtained using the total number of residents per municipality. With reference to the variable *Net foreigners* a normalization has been applied to avoid negative or zero values.

Some of the explanatory variables have been categorized for methodological reasons (for more technical details, see Rogge et al., 2017, and the references therein). Specifically, *Income per capita* and *Financial debt per capita* has been divided into deciles, *Population growth* into quartiles and *Ideological Complexion of the local Government (ICG)* has been split in three parts.

Out of 308 Flemish municipalities, one municipality has not been included in the analysis because of the lack of data.

Table A.1: Downloaded variables used for the empirical application.

DOMEIN	*SUBDOMEIN*	*INDICATOR*	INDICATOR
Arbeidsmarkt	*Werkloosheid*	*Werkloosheidsgraad (15-64 jaar) (Steunpunt Werk)*	Unemployment rate (15-64) (Centre for Work)
Criminaliteit	*Geregistreerde misdrijven*	*Diefstallen en afpersingen - per 1.000 inwoners*	Thefts and extortions - per 1,000 inhabitants
Criminaliteit	*Geregistreerde misdrijven*	*Misdrijven tegen eigendom - per 1.000 inwoners*	Crimes against property - per 1,000 inhabitants
Criminaliteit	*Geregistreerde misdrijven*	*Misdrijven tegen lichamelijke integriteit - per 1.000 inwoners*	Crimes against physical integrity - per 1,000 inhabitants
Cultuur	*Algemeen*	*Cultuurevenementen: aantal per 1.000 inwoners*	Culture Events: number per 1,000 inhabitants
Demografie	*Structuur bevolking*	*Inwoners - totaal*	Residents - total
Demografie	*Structuur bevolking*	*Inwoners 65 jaar en ouder - aandeel*	Residents age 65 and older - share
Demografie	*Structuur bevolking*	*Ouderen 80 jaar en ouder - aantal*	Older people aged 80 and over - Number
Demografie	*Huishoudens*	*Private huishoudens - totaal aantal*	Private households - total
Demografie	*Migraties*	*Saldo internationale migraties van vreemdelingen*	Net international migration of foreigners
Economie en innovatie	*Macro-economie*	*Belastbaar inkomen per inwoner*	Taxable income per capita
Energie	*Energieverbruik*	*Energieverbruik per inwoner door verwarming huishoudens*	Energy consumption per capita by household heating
Financin & Bestuur	*Gemeentebelastingen*	*Totale belastingontvangsten per inwoner*	Total tax revenue per capita
Financin & Bestuur	*Financile schuld gemeenten*	*Financile schuld per inwoner*	Financial debt per capita
Inburgering & Integratie	*Aanwezigheid*	*Vreemdelingen - aandeel t.o.v. totale bevolking*	Foreigners - share relative to total population
Milieu en natuur	*Afval*	*Restafval in kg per inwoner*	Waste in kg per capita
Mobiliteit	*Verkeersveiligheid*	*Ongevallen - aantal*	Accidents - number
Onderwijs en vorming	*Kleuteronderwijs*	*Totaal kleuteronderwijs - aantal leerlingen (naar vestigingplaats)*	Total kindergarten - number of students (by domicile)
Onderwijs en vorming	*Lager onderwijs*	*Totaal lager onderwijs - aantal leerlingen (naar vestigingplaats)*	Total primary education - number of students (by domicile)
Ruimtelijke ontwikkeling	*Oppervlakte*	*Bebouwde oppervlakte*	Built up area

Note: The first three columns show the path followed to get the chosen variables. Accordingly, the original Dutch names are reported. The last column presents the English version as translated by the authors.

A.2 Municipal-specific weights

In the following, the specific lower and upper bound weights for each Flemish municipality and for each municipal function are listed.

Table A.2: Lower bound municipal-specific weights

Municipality	Administration	Culture	Care services	Education	Housing	Local mobility	Security	Environment
Aalst	0.09	0.06	0.08	0.05	0.03	0.05	0.09	0.06
Aalter	0.11	0.08	0.06	0.04	0.03	0.06	0.06	0.05
Aarschot	0.04	0.05	0.04	0.17	0.03	0.07	0.05	0.03
Aartselaar	0.08	0.10	0.06	0.08	0.01	0.08	0.06	0.04
Affligem	0.14	0.05	0.06	0.00	0.02	0.09	0.06	0.06
Alken	0.11	0.09	0.05	0.06	0.01	0.07	0.04	0.05
Alveringem	0.14	0.06	0.03	0.03	0.02	0.12	0.04	0.04
Antwerpen	0.11	0.06	0.08	0.02	0.04	0.02	0.10	0.04
Anzegem	0.07	0.07	0.06	0.06	0.04	0.08	0.04	0.07
Ardooie	0.07	0.09	0.06	0.07	0.03	0.09	0.05	0.05
Arendonk	0.09	0.05	0.06	0.11	0.01	0.07	0.05	0.07
As	0.10	0.06	0.07	0.06	0.02	0.08	0.04	0.04
Asse	0.08	0.07	0.05	0.10	0.04	0.04	0.08	0.03
Assenede	0.12	0.05	0.04	0.02	0.03	0.09	0.06	0.04
Avelgem	0.07	0.10	0.06	0.01	0.03	0.10	0.05	0.05
Baarle-Hertog	0.17	0.08	0.01	0.01	0.02	0.12	0.05	0.05
Balen	0.07	0.09	0.02	0.10	0.03	0.08	0.05	0.06
Beernem	0.08	0.10	0.06	0.03	0.01	0.13	0.06	0.04
Beerse	0.06	0.05	0.06	0.12	0.02	0.07	0.04	0.07
Beersel	0.08	0.06	0.06	0.07	0.03	0.11	0.06	0.03
Begijnendijk	0.10	0.05	0.06	0.11	0.02	0.05	0.04	0.05
Bekkevoort	0.13	0.05	0.05	0.06	0.02	0.08	0.04	0.07
Beringen	0.09	0.06	0.08	0.06	0.02	0.05	0.04	0.08
Berlaar	0.10	0.04	0.06	0.09	0.03	0.08	0.07	0.05
Berlare	0.09	0.08	0.09	0.02	0.02	0.08	0.05	0.06
Bertem	0.13	0.07	0.03	0.05	0.01	0.10	0.05	0.06
Bever	0.14	0.01	0.05	0.08	0.01	0.12	0.04	0.04
Beveren	0.07	0.07	0.06	0.11	0.04	0.04	0.04	0.04
Bierbeek	0.09	0.09	0.06	0.05	0.01	0.07	0.05	0.07
Bilzen	0.08	0.08	0.04	0.06	0.04	0.06	0.04	0.07
Blankenberge	0.09	0.10	0.06	0.00	0.04	0.02	0.09	0.03
Bocholt	0.09	0.16	0.06	0.00	0.01	0.10	0.03	0.03
Boechout	0.11	0.05	0.06	0.04	0.02	0.12	0.05	0.04
Bonheiden	0.09	0.08	0.05	0.06	0.03	0.07	0.06	0.06
Boom	0.10	0.06	0.10	0.02	0.03	0.04	0.08	0.04
Boortmeerbeek	0.09	0.04	0.07	0.07	0.01	0.09	0.05	0.09
Borgloon	0.13	0.08	0.06	0.00	0.03	0.09	0.04	0.04
Bornem	0.08	0.09	0.05	0.03	0.04	0.09	0.06	0.06
Borsbeek	0.11	0.05	0.07	0.11	0.01	0.06	0.06	0.03
Boutersem	0.10	0.05	0.05	0.05	0.01	0.08	0.05	0.10
Brakel	0.10	0.08	0.06	0.03	0.01	0.09	0.06	0.06
Brasschaat	0.08	0.05	0.06	0.14	0.02	0.07	0.05	0.03
Brecht	0.08	0.06	0.05	0.05	0.04	0.11	0.06	0.04
Bredene	0.10	0.07	0.09	0.00	0.03	0.09	0.08	0.02

Bree	0.07	0.09	0.06	0.01	0.02	0.11	0.06	0.04
Brugge	0.10	0.08	0.07	0.04	0.05	0.02	0.08	0.04
Buggenhout	0.11	0.08	0.07	0.05	0.02	0.04	0.05	0.05
Damme	0.08	0.10	0.05	0.03	0.02	0.10	0.05	0.06
De Haan	0.06	0.10	0.05	0.02	0.03	0.12	0.09	0.02
De Panne	0.07	0.07	0.07	0.02	0.04	0.09	0.08	0.04
De Pinte	0.08	0.07	0.07	0.06	0.02	0.08	0.05	0.06
Deerlijk	0.25	0.04	0.03	0.02	0.04	0.06	0.04	0.03
Deinze	0.08	0.08	0.07	0.04	0.06	0.08	0.05	0.04
Denderleeuw	0.12	0.05	0.08	0.06	0.04	0.04	0.06	0.05
Dendermonde	0.08	0.09	0.10	0.05	0.03	0.03	0.06	0.05
Dentergem	0.08	0.09	0.04	0.10	0.02	0.07	0.05	0.05
Dessel	0.07	0.06	0.07	0.06	0.04	0.08	0.05	0.05
Destelbergen	0.08	0.04	0.08	0.08	0.03	0.08	0.06	0.06
Diepenbeek	0.10	0.05	0.07	0.07	0.02	0.07	0.04	0.05
Diest	0.07	0.07	0.07	0.05	0.03	0.08	0.06	0.06
Diksmuide	0.07	0.09	0.05	0.05	0.06	0.07	0.05	0.06
Dilbeek	0.09	0.08	0.07	0.08	0.03	0.04	0.06	0.06
Dilsen-Stokkem	0.07	0.05	0.07	0.14	0.01	0.07	0.04	0.04
Drogenbos	0.12	0.03	0.04	0.14	0.02	0.04	0.07	0.02
Duffel	0.06	0.07	0.03	0.12	0.06	0.07	0.05	0.04
Edegem	0.15	0.05	0.06	0.05	0.03	0.04	0.07	0.03
Eeklo	0.09	0.08	0.08	0.02	0.03	0.06	0.08	0.06
Erpe-Mere	0.10	0.05	0.05	0.03	0.02	0.09	0.06	0.08
Essen	0.09	0.05	0.07	0.04	0.04	0.08	0.06	0.05
Evergem	0.09	0.08	0.08	0.04	0.02	0.07	0.05	0.04
Galmaarden	0.08	0.05	0.05	0.11	0.01	0.08	0.05	0.06
Gavere	0.09	0.05	0.03	0.10	0.04	0.08	0.07	0.04
Geel	0.13	0.07	0.05	0.06	0.01	0.04	0.08	0.05
Geetbets	0.10	0.06	0.06	0.00	0.02	0.12	0.05	0.06
Genk	0.09	0.08	0.09	0.02	0.03	0.04	0.07	0.03
Gent	0.10	0.06	0.10	0.07	0.03	0.03	0.07	0.03
Geraardsbergen	0.08	0.06	0.10	0.02	0.03	0.06	0.06	0.06
Gingelom	0.09	0.04	0.03	0.15	0.01	0.09	0.04	0.04
Gistel	0.09	0.08	0.06	0.03	0.08	0.07	0.06	0.03
Glabbeek	0.10	0.06	0.06	0.06	0.01	0.12	0.04	0.04
Gooik	0.09	0.05	0.05	0.09	0.01	0.09	0.05	0.06
Grimbergen	0.10	0.10	0.05	0.07	0.02	0.05	0.06	0.03
Grobbendonk	0.08	0.04	0.06	0.06	0.04	0.06	0.06	0.08
Haacht	0.11	0.05	0.06	0.05	0.05	0.05	0.06	0.06
Haaltert	0.11	0.06	0.07	0.01	0.02	0.09	0.07	0.05
Halen	0.09	0.09	0.05	0.05	0.02	0.10	0.05	0.03
Halle	0.08	0.06	0.09	0.05	0.02	0.06	0.09	0.04
Ham	0.10	0.08	0.06	0.00	0.01	0.08	0.04	0.07
Hamme	0.13	0.07	0.06	0.06	0.01	0.03	0.06	0.06
Hamont-Achel	0.08	0.12	0.06	0.00	0.01	0.11	0.04	0.06
Harelbeke	0.08	0.07	0.05	0.10	0.03	0.06	0.05	0.06
Hasselt	0.08	0.07	0.08	0.05	0.02	0.07	0.08	0.03
Hechtel-Eksel	0.10	0.08	0.08	0.05	0.01	0.08	0.03	0.04
Heers	0.12	0.06	0.05	0.01	0.01	0.13	0.06	0.04
Heist-op-den-Berg	0.06	0.06	0.06	0.04	0.01	0.07	0.12	0.05
Hemiksem	0.17	0.04	0.04	0.14	0.02	0.02	0.05	0.03
Herent	0.17	0.05	0.06	0.01	0.04	0.06	0.04	0.05

Herentals	0.11	0.08	0.06	0.04	0.05	0.05	0.06	0.05
Herenthout	0.09	0.05	0.06	0.10	0.03	0.08	0.04	0.04
Herk-de-Stad	0.08	0.11	0.10	0.00	0.01	0.10	0.05	0.04
Herne	0.10	0.06	0.05	0.03	0.01	0.09	0.07	0.08
Herselt	0.11	0.04	0.05	0.00	0.03	0.11	0.04	0.04
Herstappe	0.38	0.00	0.00	0.00	0.02	0.02	0.05	0.03
Herzele	0.09	0.07	0.09	0.05	0.02	0.08	0.06	0.04
Heusden-Zolder	0.08	0.07	0.07	0.09	0.02	0.05	0.06	0.05
Heuvelland	0.09	0.08	0.04	0.03	0.03	0.10	0.05	0.06
Hoegaarden	0.11	0.06	0.04	0.10	0.02	0.08	0.04	0.04
Hoeilaart	0.06	0.10	0.05	0.00	0.07	0.08	0.04	0.04
Hoeselt	0.09	0.08	0.06	0.02	0.01	0.11	0.05	0.05
Holsbeek	0.08	0.07	0.02	0.08	0.01	0.15	0.05	0.05
Hooglede	0.09	0.07	0.08	0.02	0.04	0.09	0.05	0.06
Hoogstraten	0.07	0.05	0.06	0.12	0.02	0.08	0.06	0.04
Horebeke	0.15	0.07	0.03	0.00	0.02	0.09	0.06	0.07
Houthalen-Helchteren	0.22	0.05	0.06	0.06	0.01	0.02	0.06	0.03
Houthulst	0.08	0.09	0.06	0.03	0.02	0.11	0.04	0.06
Hove	0.13	0.07	0.08	0.07	0.02	0.08	0.05	0.02
Huldenberg	0.09	0.05	0.06	0.11	0.02	0.06	0.05	0.05
Hulshout	0.09	0.02	0.09	0.08	0.03	0.10	0.04	0.06
Ichtegem	0.09	0.09	0.06	0.06	0.04	0.06	0.04	0.07
Ieper	0.07	0.14	0.04	0.03	0.02	0.06	0.05	0.04
Ingelmunster	0.10	0.07	0.06	0.06	0.02	0.10	0.05	0.05
Izegem	0.07	0.10	0.06	0.03	0.03	0.08	0.06	0.07
Jabbeke	0.07	0.11	0.06	0.00	0.06	0.08	0.04	0.04
Kalmthout	0.04	0.04	0.05	0.16	0.02	0.09	0.05	0.04
Kampenhout	0.11	0.05	0.08	0.07	0.01	0.06	0.05	0.06
Kapelle-op-den-Bos	0.10	0.07	0.06	0.06	0.02	0.07	0.06	0.07
Kapellen	0.09	0.09	0.07	0.02	0.04	0.06	0.07	0.04
Kaprijke	0.12	0.06	0.08	0.01	0.02	0.07	0.06	0.07
Kasterlee	0.14	0.06	0.06	0.07	0.02	0.02	0.06	0.06
Keerbergen	0.09	0.04	0.06	0.06	0.02	0.07	0.06	0.09
Kinrooi	0.09	0.07	0.06	0.07	0.01	0.12	0.03	0.04
Kluisbergen	0.07	0.07	0.03	0.08	0.06	0.08	0.05	0.06
Knesselare	0.11	0.10	0.07	0.01	0.03	0.05	0.05	0.06
Knokke-Heist	0.06	0.11	0.06	0.04	0.06	0.03	0.09	0.03
Koekelare	0.08	0.11	0.05	0.00	0.03	0.12	0.04	0.07
Koksijde	0.09	0.13	0.04	0.04	0.03	0.04	0.06	0.04
Kontich	0.09	0.09	0.07	0.04	0.05	0.06	0.07	0.04
Kortemark	0.06	0.08	0.05	0.05	0.03	0.08	0.04	0.08
Kortenaken	0.07	0.06	0.08	0.01	0.02	0.16	0.04	0.05
Kortenberg	0.09	0.06	0.04	0.07	0.04	0.07	0.06	0.06
Kortessem	0.09	0.10	0.04	0.08	0.01	0.08	0.04	0.04
Kortrijk	0.08	0.10	0.08	0.02	0.05	0.05	0.07	0.04
Kraainem	0.10	0.02	0.05	0.07	0.03	0.08	0.09	0.04
Kruibeke	0.09	0.06	0.10	0.03	0.02	0.08	0.06	0.04
Kruishoutem	0.10	0.07	0.03	0.07	0.05	0.10	0.05	0.03
Kuurne	0.06	0.08	0.05	0.08	0.06	0.05	0.06	0.06
Laakdal	0.09	0.06	0.04	0.10	0.01	0.09	0.05	0.05
Laarne	0.06	0.07	0.07	0.05	0.03	0.10	0.05	0.05
Lanaken	0.09	0.09	0.05	0.06	0.01	0.08	0.06	0.06
Landen	0.10	0.06	0.09	0.04	0.01	0.08	0.07	0.04

Langemark-Poelkapelle	0.11	0.09	0.05	0.03	0.03	0.10	0.03	0.05
Lebbeke	0.07	0.10	0.05	0.03	0.02	0.08	0.06	0.04
Lede	0.08	0.08	0.07	0.02	0.03	0.09	0.06	0.07
Ledegem	0.09	0.09	0.05	0.01	0.06	0.06	0.07	0.05
Lendelede	0.08	0.12	0.04	0.01	0.03	0.09	0.05	0.08
Lennik	0.06	0.04	0.03	0.06	0.03	0.12	0.10	0.04
Leopoldsburg	0.09	0.12	0.09	0.00	0.01	0.09	0.06	0.04
Leuven	0.09	0.09	0.09	0.03	0.04	0.04	0.09	0.03
Lichtervelde	0.07	0.08	0.05	0.07	0.01	0.07	0.05	0.08
Liedekerke	0.10	0.07	0.05	0.05	0.04	0.07	0.06	0.07
Lier	0.11	0.09	0.07	0.01	0.04	0.05	0.08	0.05
Lierde	0.10	0.12	0.03	0.01	0.03	0.08	0.05	0.07
Lille	0.08	0.04	0.03	0.07	0.04	0.06	0.05	0.10
Linkebeek	0.08	0.02	0.06	0.14	0.02	0.09	0.06	0.02
Lint	0.10	0.07	0.07	0.05	0.03	0.09	0.05	0.04
Linter	0.10	0.03	0.06	0.07	0.02	0.11	0.05	0.04
Lo-Reninge	0.15	0.09	0.02	0.00	0.04	0.11	0.04	0.04
Lochristi	0.06	0.06	0.05	0.09	0.03	0.09	0.04	0.07
Lokeren	0.08	0.06	0.06	0.08	0.04	0.04	0.07	0.07
Lommel	0.06	0.09	0.05	0.03	0.01	0.07	0.07	0.03
Londerzeel	0.05	0.06	0.07	0.04	0.03	0.10	0.09	0.04
Lovendegem	0.08	0.05	0.06	0.07	0.02	0.07	0.05	0.08
Lubbeek	0.09	0.06	0.07	0.10	0.01	0.06	0.05	0.06
Lummen	0.08	0.10	0.09	0.02	0.01	0.09	0.05	0.06
Maarkedal	0.11	0.05	0.06	0.08	0.03	0.07	0.05	0.05
Maaseik	0.07	0.12	0.08	0.02	0.01	0.09	0.05	0.03
Maasmechelen	0.08	0.08	0.12	0.03	0.01	0.08	0.06	0.04
Machelen	0.08	0.06	0.12	0.08	0.04	0.03	0.06	0.02
Maldegem	0.09	0.08	0.09	0.03	0.04	0.07	0.06	0.05
Malle	0.08	0.07	0.08	0.03	0.04	0.07	0.08	0.04
Mechelen	0.09	0.07	0.07	0.02	0.05	0.02	0.10	0.05
Meerhout	0.10	0.03	0.09	0.05	0.07	0.06	0.05	0.05
Meeuwen-Gruitrode	0.12	0.11	0.04	0.01	0.02	0.12	0.02	0.05
Meise	0.10	0.07	0.04	0.13	0.02	0.05	0.04	0.04
Melle	0.09	0.07	0.06	0.05	0.03	0.08	0.07	0.05
Menen	0.08	0.08	0.08	0.05	0.03	0.05	0.08	0.03
Merchtem	0.06	0.04	0.05	0.20	0.01	0.05	0.03	0.03
Merelbeke	0.07	0.08	0.07	0.04	0.04	0.08	0.05	0.06
Merksplas	0.06	0.12	0.06	0.06	0.04	0.08	0.03	0.04
Mesen	0.17	0.09	0.03	0.00	0.02	0.08	0.05	0.04
Meulebeke	0.08	0.10	0.06	0.00	0.05	0.09	0.06	0.05
Middelkerke	0.05	0.09	0.05	0.03	0.06	0.10	0.05	0.05
Moerbeke	0.07	0.05	0.08	0.08	0.02	0.07	0.04	0.07
Mol	0.14	0.06	0.08	0.03	0.03	0.03	0.07	0.05
Moorslede	0.07	0.13	0.03	0.05	0.02	0.09	0.04	0.06
Mortsel	0.09	0.06	0.06	0.13	0.04	0.03	0.05	0.02
Nazareth	0.07	0.11	0.08	0.06	0.03	0.07	0.05	0.04
Neerpelt	0.09	0.07	0.05	0.00	0.02	0.15	0.04	0.07
Nevele	0.08	0.06	0.10	0.06	0.04	0.07	0.06	0.03
Niel	0.15	0.05	0.04	0.08	0.01	0.09	0.05	0.04
Nieuwerkerken	0.12	0.05	0.06	0.02	0.01	0.09	0.05	0.05
Nieuwpoort	0.09	0.08	0.05	0.01	0.05	0.04	0.08	0.05
Nijlen	0.07	0.04	0.03	0.16	0.01	0.08	0.05	0.04

Ninove	0.05	0.05	0.06	0.08	0.03	0.10	0.06	0.07
Olen	0.11	0.05	0.06	0.05	0.02	0.07	0.05	0.08
Oostende	0.09	0.06	0.06	0.04	0.05	0.03	0.10	0.04
Oosterzele	0.08	0.05	0.05	0.05	0.02	0.09	0.04	0.07
Oostkamp	0.10	0.12	0.06	0.00	0.04	0.05	0.06	0.06
Oostrozebeke	0.12	0.08	0.08	0.00	0.05	0.07	0.05	0.04
Opglabbeek	0.15	0.10	0.05	0.00	0.02	0.07	0.04	0.05
Opwijk	0.07	0.05	0.06	0.15	0.02	0.05	0.03	0.04
Oud-Heverlee	0.07	0.06	0.05	0.16	0.01	0.07	0.04	0.04
Oud-Turnhout	0.10	0.05	0.05	0.04	0.03	0.06	0.07	0.08
Oudenaarde	0.06	0.10	0.06	0.01	0.04	0.06	0.08	0.07
Oudenburg	0.07	0.11	0.06	0.00	0.03	0.11	0.05	0.06
Overijse	0.10	0.06	0.07	0.08	0.02	0.05	0.06	0.04
Overpelt	0.11	0.08	0.11	0.01	0.06	0.08	0.03	0.03
Peer	0.14	0.06	0.08	0.00	0.02	0.11	0.04	0.03
Pepingen	0.09	0.07	0.05	0.00	0.02	0.12	0.06	0.09
Pittem	0.08	0.07	0.08	0.00	0.05	0.09	0.04	0.07
Poperinge	0.06	0.12	0.04	0.03	0.05	0.09	0.05	0.05
Putte	0.07	0.05	0.08	0.00	0.03	0.08	0.05	0.09
Puurs	0.08	0.07	0.06	0.00	0.06	0.10	0.06	0.05
Ranst	0.07	0.05	0.04	0.07	0.05	0.11	0.07	0.05
Ravels	0.09	0.06	0.04	0.09	0.01	0.11	0.04	0.07
Retie	0.08	0.04	0.06	0.07	0.02	0.08	0.04	0.09
Riemst	0.08	0.07	0.03	0.07	0.03	0.11	0.04	0.06
Rijkevorsel	0.10	0.04	0.05	0.03	0.02	0.17	0.05	0.05
Roeselare	0.07	0.08	0.09	0.06	0.04	0.06	0.07	0.02
Ronse	0.07	0.07	0.11	0.05	0.05	0.05	0.06	0.04
Roosdaal	0.10	0.05	0.05	0.09	0.03	0.07	0.06	0.05
Rotselaar	0.08	0.05	0.05	0.06	0.02	0.07	0.04	0.07
Ruiselede	0.11	0.10	0.04	0.01	0.02	0.09	0.05	0.08
Rumst	0.10	0.05	0.04	0.08	0.06	0.08	0.05	0.04
Schelle	0.09	0.04	0.07	0.09	0.04	0.08	0.05	0.04
Scherpenheuvel-Zichem	0.05	0.10	0.06	0.05	0.03	0.10	0.05	0.03
Schilde	0.09	0.06	0.05	0.02	0.03	0.10	0.07	0.07
Schoten	0.09	0.08	0.09	0.01	0.03	0.06	0.08	0.05
Sint-Amands	0.09	0.08	0.06	0.00	0.02	0.12	0.05	0.07
Sint-Genesius-Rode	0.07	0.03	0.08	0.07	0.04	0.07	0.08	0.06
Sint-Gillis-Waas	0.09	0.07	0.06	0.06	0.03	0.08	0.05	0.06
Sint-Katelijne-Waver	0.09	0.06	0.07	0.06	0.06	0.05	0.05	0.05
Sint-Laureins	0.07	0.06	0.04	0.07	0.02	0.09	0.06	0.10
Sint-Lievens-Houtem	0.09	0.08	0.04	0.03	0.02	0.05	0.07	0.07
Sint-Martens-Latem	0.07	0.08	0.04	0.10	0.02	0.08	0.06	0.05
Sint-Niklaas	0.09	0.07	0.11	0.05	0.02	0.02	0.07	0.06
Sint-Pieters-Leeuw	0.07	0.07	0.08	0.06	0.04	0.04	0.07	0.05
Sint-Truiden	0.08	0.10	0.07	0.03	0.04	0.05	0.06	0.04
Spiere-Helkijn	0.09	0.09	0.04	0.00	0.08	0.12	0.03	0.04
Stabroek	0.13	0.05	0.07	0.06	0.02	0.07	0.06	0.05
Staden	0.08	0.06	0.07	0.06	0.03	0.09	0.04	0.06
Steenokkerzeel	0.10	0.06	0.04	0.09	0.02	0.04	0.05	0.05
Stekene	0.10	0.06	0.06	0.04	0.05	0.04	0.05	0.07
Temse	0.11	0.07	0.07	0.01	0.04	0.04	0.06	0.08
Ternat	0.09	0.11	0.05	0.05	0.04	0.07	0.06	0.04
Tervuren	0.10	0.05	0.05	0.07	0.01	0.09	0.06	0.04

Tessenderlo	0.08	0.10	0.07	0.03	0.03	0.08	0.06	0.06
Tielt	0.08	0.12	0.08	0.01	0.02	0.07	0.04	0.05
Tielt-Winge	0.11	0.08	0.10	0.00	0.01	0.09	0.04	0.05
Tienen	0.07	0.10	0.07	0.02	0.03	0.07	0.08	0.05
Tongeren	0.16	0.08	0.02	0.05	0.01	0.06	0.05	0.03
Torhout	0.07	0.09	0.05	0.01	0.07	0.07	0.06	0.06
Tremelo	0.13	0.04	0.09	0.00	0.01	0.10	0.06	0.06
Turnhout	0.07	0.08	0.08	0.08	0.05	0.04	0.07	0.03
Veurne	0.10	0.08	0.07	0.04	0.04	0.07	0.05	0.05
Vilvoorde	0.11	0.05	0.07	0.10	0.04	0.02	0.09	0.03
Vleteren	0.10	0.11	0.02	0.00	0.02	0.12	0.05	0.06
Voeren	0.13	0.03	0.06	0.00	0.02	0.10	0.04	0.11
Vorselaar	0.12	0.05	0.06	0.06	0.04	0.07	0.05	0.05
Vosselaar	0.07	0.04	0.04	0.09	0.05	0.06	0.04	0.10
Waarschoot	0.10	0.09	0.05	0.01	0.03	0.10	0.05	0.03
Waasmunster	0.08	0.05	0.07	0.04	0.02	0.07	0.05	0.08
Wachtebeke	0.11	0.09	0.11	0.00	0.02	0.06	0.05	0.08
Waregem	0.07	0.10	0.07	0.08	0.02	0.05	0.04	0.07
Wellen	0.11	0.08	0.04	0.00	0.01	0.13	0.05	0.05
Wemmel	0.08	0.02	0.10	0.13	0.03	0.04	0.05	0.04
Wervik	0.07	0.11	0.07	0.00	0.07	0.09	0.05	0.02
Westerlo	0.08	0.05	0.06	0.07	0.03	0.08	0.06	0.07
Wetteren	0.09	0.09	0.06	0.04	0.04	0.05	0.06	0.05
Wevelgem	0.08	0.10	0.08	0.04	0.06	0.05	0.06	0.03
Wezembeek-Oppem	0.08	0.03	0.05	0.08	0.02	0.10	0.09	0.05
Wichelen	0.13	0.04	0.08	0.07	0.02	0.05	0.05	0.05
Wielsbeke	0.11	0.10	0.06	0.00	0.03	0.10	0.05	0.04
Wijnegem	0.08	0.08	0.08	0.11	0.03	0.04	0.06	0.02
Willebroek	0.12	0.06	0.10	0.00	0.03	0.07	0.07	0.04
Wingene	0.07	0.14	0.04	0.01	0.02	0.10	0.04	0.08
Wommelgem	0.10	0.07	0.05	0.07	0.04	0.08	0.06	0.04
Wortegem-Petegem	0.09	0.11	0.04	0.07	0.02	0.09	0.03	0.05
Wuustwezel	0.05	0.04	0.07	0.08	0.03	0.12	0.07	0.04
Zandhoven	0.08	0.06	0.03	0.09	0.03	0.08	0.06	0.04
Zaventem	0.10	0.05	0.06	0.06	0.05	0.03	0.12	0.03
Zedelgem	0.09	0.10	0.08	0.00	0.04	0.07	0.06	0.06
Zele	0.06	0.04	0.08	0.05	0.03	0.09	0.07	0.06
Zelzate	0.07	0.05	0.10	0.05	0.02	0.06	0.08	0.06
Zemst	0.09	0.04	0.04	0.08	0.04	0.07	0.05	0.08
Zingem	0.14	0.08	0.03	0.11	0.03	0.04	0.04	0.05
Zoersel	0.19	0.05	0.01	0.11	0.03	0.04	0.05	0.03
Zomergem	0.08	0.10	0.06	0.00	0.03	0.10	0.06	0.06
Zonhoven	0.12	0.09	0.08	0.00	0.01	0.09	0.06	0.04
Zonnebeke	0.08	0.13	0.04	0.05	0.05	0.08	0.04	0.03
Zottegem	0.18	0.06	0.06	0.07	0.02	0.02	0.07	0.04
Zoutleeuw	0.08	0.05	0.07	0.02	0.02	0.13	0.05	0.06
Zuienkerke	0.13	0.05	0.02	0.06	0.02	0.11	0.06	0.05
Zulte	0.10	0.08	0.03	0.04	0.04	0.11	0.06	0.04
Zutendaal	0.09	0.07	0.04	0.00	0.01	0.15	0.05	0.05
Zwalm	0.10	0.06	0.06	0.00	0.06	0.09	0.05	0.08
Zwevegem	0.08	0.09	0.07	0.08	0.03	0.08	0.04	0.03
Zwijndrecht	0.10	0.06	0.11	0.00	0.06	0.06	0.07	0.04

Table A.3: Upper bound municipal-specific weights

Municipality	Administration	Culture	Care services	Education	Housing	Local mobility	Security	Environment
Aalst	0.27	0.18	0.23	0.15	0.08	0.14	0.26	0.17
Aalter	0.33	0.23	0.18	0.11	0.09	0.18	0.18	0.15
Aarschot	0.12	0.15	0.11	0.51	0.09	0.21	0.14	0.09
Aartselaar	0.23	0.29	0.17	0.24	0.03	0.23	0.18	0.12
Affligem	0.41	0.15	0.18	0.00	0.05	0.26	0.17	0.17
Alken	0.32	0.26	0.15	0.18	0.02	0.21	0.11	0.15
Alveringem	0.42	0.18	0.09	0.09	0.05	0.36	0.11	0.12
Antwerpen	0.33	0.18	0.24	0.06	0.11	0.06	0.30	0.11
Anzegem	0.20	0.20	0.18	0.17	0.12	0.24	0.12	0.20
Ardooie	0.20	0.26	0.18	0.21	0.09	0.26	0.14	0.14
Arendonk	0.27	0.15	0.17	0.32	0.03	0.20	0.14	0.21
As	0.30	0.18	0.20	0.17	0.05	0.23	0.12	0.11
Asse	0.23	0.21	0.15	0.29	0.12	0.11	0.24	0.09
Assenede	0.35	0.15	0.12	0.06	0.08	0.26	0.17	0.11
Avelgem	0.20	0.29	0.17	0.02	0.08	0.29	0.15	0.15
Baarle-Hertog	0.51	0.23	0.03	0.03	0.06	0.35	0.14	0.14
Balen	0.21	0.26	0.06	0.30	0.08	0.24	0.15	0.17
Beernem	0.24	0.29	0.18	0.08	0.03	0.39	0.17	0.12
Beerse	0.18	0.15	0.18	0.35	0.06	0.20	0.12	0.21
Beersel	0.24	0.18	0.17	0.21	0.08	0.32	0.17	0.09
Begijnendijk	0.29	0.15	0.17	0.32	0.06	0.15	0.12	0.14
Bekkevoort	0.39	0.14	0.15	0.17	0.05	0.23	0.11	0.20
Beringen	0.27	0.18	0.23	0.18	0.06	0.14	0.11	0.23
Berlaar	0.29	0.11	0.17	0.26	0.09	0.23	0.20	0.14
Berlare	0.27	0.24	0.26	0.05	0.05	0.23	0.15	0.18
Bertem	0.38	0.20	0.09	0.15	0.03	0.29	0.14	0.17
Bever	0.41	0.03	0.14	0.23	0.03	0.35	0.12	0.12
Beveren	0.20	0.21	0.18	0.32	0.11	0.11	0.12	0.11
Bierbeek	0.27	0.27	0.17	0.15	0.03	0.20	0.15	0.21
Bilzen	0.24	0.24	0.12	0.17	0.11	0.17	0.11	0.20
Blankenberge	0.27	0.30	0.18	0.00	0.12	0.05	0.27	0.09
Bocholt	0.27	0.47	0.18	0.00	0.03	0.29	0.09	0.09
Boechout	0.33	0.15	0.17	0.12	0.06	0.35	0.14	0.12
Bonheiden	0.27	0.24	0.14	0.17	0.08	0.21	0.18	0.17
Boom	0.30	0.17	0.30	0.06	0.09	0.11	0.24	0.12
Boortmeerbeek	0.26	0.12	0.20	0.21	0.03	0.26	0.14	0.27
Borgloon	0.38	0.23	0.18	0.00	0.08	0.27	0.12	0.11
Bornem	0.23	0.26	0.15	0.08	0.11	0.27	0.17	0.17
Borsbeek	0.32	0.14	0.21	0.33	0.03	0.17	0.17	0.09
Boutersem	0.30	0.15	0.15	0.14	0.03	0.24	0.14	0.29
Brakel	0.29	0.24	0.18	0.08	0.03	0.27	0.17	0.18
Brasschaat	0.23	0.14	0.18	0.42	0.06	0.21	0.14	0.08
Brecht	0.24	0.17	0.14	0.15	0.11	0.33	0.17	0.12
Bredene	0.29	0.21	0.27	0.00	0.09	0.26	0.23	0.06
Bree	0.20	0.26	0.17	0.02	0.06	0.33	0.17	0.12
Brugge	0.30	0.24	0.21	0.12	0.14	0.06	0.24	0.11
Buggenhout	0.33	0.24	0.21	0.15	0.06	0.12	0.14	0.14
Damme	0.24	0.30	0.15	0.08	0.06	0.29	0.14	0.18
De Haan	0.18	0.29	0.14	0.05	0.09	0.36	0.26	0.06

De Panne	0.20	0.20	0.21	0.06	0.11	0.26	0.23	0.12
De Pinte	0.23	0.21	0.21	0.18	0.06	0.23	0.14	0.17
Deerlijk	0.74	0.11	0.09	0.05	0.11	0.17	0.11	0.08
Deinze	0.23	0.23	0.20	0.12	0.17	0.23	0.15	0.12
Denderleeuw	0.35	0.15	0.23	0.17	0.12	0.12	0.17	0.14
Dendermonde	0.23	0.26	0.29	0.14	0.08	0.09	0.18	0.15
Dentergem	0.23	0.26	0.11	0.29	0.05	0.20	0.15	0.15
Dessel	0.21	0.18	0.20	0.18	0.11	0.23	0.14	0.15
Destelbergen	0.23	0.12	0.23	0.23	0.09	0.23	0.17	0.17
Diepenbeek	0.30	0.15	0.20	0.20	0.05	0.20	0.12	0.14
Diest	0.20	0.21	0.20	0.14	0.08	0.24	0.18	0.18
Diksmuide	0.20	0.27	0.14	0.14	0.18	0.20	0.14	0.18
Dilbeek	0.26	0.24	0.20	0.23	0.08	0.12	0.17	0.18
Dilsen-Stokkem	0.21	0.15	0.20	0.41	0.02	0.20	0.12	0.11
Drogenbos	0.36	0.09	0.11	0.41	0.06	0.12	0.20	0.06
Duffel	0.17	0.20	0.09	0.36	0.17	0.21	0.15	0.11
Edegem	0.44	0.15	0.17	0.15	0.08	0.12	0.21	0.09
Eeklo	0.26	0.24	0.24	0.05	0.09	0.17	0.23	0.17
Erpe-Mere	0.30	0.15	0.15	0.09	0.05	0.27	0.18	0.24
Essen	0.27	0.15	0.20	0.12	0.11	0.23	0.17	0.14
Evergem	0.26	0.24	0.23	0.12	0.06	0.21	0.15	0.11
Galmaarden	0.24	0.15	0.15	0.33	0.03	0.23	0.14	0.18
Gavere	0.27	0.14	0.09	0.29	0.11	0.23	0.20	0.12
Geel	0.39	0.21	0.14	0.17	0.03	0.12	0.23	0.14
Geetbets	0.30	0.18	0.18	0.00	0.05	0.36	0.14	0.18
Genk	0.26	0.23	0.26	0.05	0.09	0.12	0.21	0.09
Gent	0.29	0.18	0.29	0.21	0.08	0.08	0.21	0.09
Geraardsbergen	0.24	0.18	0.29	0.05	0.09	0.18	0.18	0.18
Gingelom	0.27	0.11	0.08	0.44	0.03	0.26	0.11	0.12
Gistel	0.26	0.24	0.17	0.08	0.24	0.21	0.17	0.09
Glabbeek	0.29	0.17	0.17	0.17	0.03	0.35	0.12	0.12
Gooik	0.27	0.15	0.15	0.26	0.02	0.26	0.15	0.18
Grimbergen	0.29	0.30	0.15	0.20	0.05	0.14	0.18	0.09
Grobbendonk	0.24	0.12	0.18	0.18	0.11	0.18	0.17	0.24
Haacht	0.33	0.15	0.18	0.14	0.14	0.15	0.17	0.17
Haaltert	0.33	0.18	0.21	0.03	0.06	0.26	0.20	0.14
Halen	0.26	0.26	0.14	0.15	0.06	0.29	0.14	0.09
Halle	0.24	0.18	0.26	0.14	0.06	0.18	0.27	0.11
Ham	0.30	0.23	0.18	0.00	0.02	0.23	0.12	0.21
Hamme	0.38	0.20	0.17	0.17	0.03	0.09	0.18	0.17
Hamont-Achel	0.23	0.35	0.18	0.00	0.02	0.33	0.11	0.17
Harelbeke	0.24	0.20	0.15	0.29	0.09	0.18	0.15	0.17
Hasselt	0.24	0.21	0.23	0.15	0.06	0.20	0.23	0.09
Hechtel-Eksel	0.29	0.24	0.23	0.15	0.03	0.24	0.09	0.11
Heers	0.35	0.18	0.15	0.03	0.02	0.38	0.17	0.12
Heist-op-den-Berg	0.17	0.17	0.17	0.11	0.03	0.21	0.35	0.15
Hemiksem	0.50	0.11	0.12	0.41	0.06	0.05	0.14	0.09
Herent	0.50	0.15	0.18	0.03	0.12	0.17	0.12	0.15
Herentals	0.32	0.24	0.18	0.11	0.15	0.15	0.17	0.14
Herenthout	0.26	0.15	0.18	0.30	0.08	0.24	0.12	0.12
Herk-de-Stad	0.23	0.33	0.29	0.00	0.03	0.30	0.15	0.11
Herne	0.30	0.18	0.15	0.09	0.03	0.26	0.20	0.23
Herselt	0.32	0.11	0.14	0.00	0.08	0.33	0.12	0.11

Herstappe	1.14	0.00	0.00	0.00	0.06	0.06	0.14	0.09
Herzele	0.26	0.21	0.26	0.15	0.05	0.23	0.17	0.12
Heusden-Zolder	0.24	0.21	0.21	0.26	0.06	0.15	0.18	0.14
Heuvelland	0.26	0.24	0.11	0.08	0.08	0.30	0.15	0.18
Hoegaarden	0.33	0.17	0.12	0.30	0.05	0.24	0.11	0.12
Hoeilaart	0.18	0.29	0.15	0.00	0.21	0.24	0.12	0.11
Hoeselt	0.27	0.24	0.18	0.05	0.02	0.33	0.15	0.15
Holsbeek	0.23	0.20	0.05	0.23	0.02	0.45	0.14	0.14
Hooglede	0.26	0.20	0.23	0.06	0.11	0.27	0.14	0.18
Hoogstraten	0.21	0.14	0.18	0.35	0.05	0.23	0.18	0.12
Horebeke	0.45	0.21	0.08	0.00	0.05	0.27	0.18	0.21
Houthalen-Helchteren	0.65	0.15	0.17	0.17	0.03	0.06	0.17	0.09
Houthulst	0.24	0.27	0.18	0.08	0.05	0.32	0.11	0.18
Hove	0.38	0.21	0.23	0.20	0.05	0.23	0.14	0.06
Huldenberg	0.26	0.14	0.17	0.32	0.05	0.18	0.14	0.15
Hulshout	0.27	0.05	0.26	0.23	0.08	0.29	0.12	0.17
Ichtegem	0.26	0.26	0.17	0.18	0.11	0.17	0.12	0.20
Ieper	0.21	0.42	0.12	0.08	0.05	0.17	0.14	0.11
Ingelmunster	0.29	0.21	0.17	0.18	0.05	0.30	0.14	0.14
Izegem	0.20	0.29	0.18	0.08	0.08	0.23	0.18	0.20
Jabbeke	0.20	0.32	0.18	0.00	0.18	0.24	0.12	0.12
Kalmthout	0.12	0.12	0.14	0.48	0.06	0.27	0.14	0.11
Kampenhout	0.32	0.14	0.23	0.21	0.03	0.17	0.14	0.18
Kapelle-op-den-Bos	0.29	0.20	0.17	0.18	0.05	0.20	0.17	0.20
Kapellen	0.26	0.26	0.21	0.05	0.11	0.18	0.21	0.11
Kaprijke	0.36	0.17	0.23	0.02	0.05	0.21	0.18	0.21
Kasterlee	0.41	0.18	0.17	0.20	0.06	0.06	0.17	0.18
Keerbergen	0.26	0.11	0.18	0.17	0.06	0.20	0.17	0.26
Kinrooi	0.26	0.20	0.18	0.21	0.02	0.35	0.08	0.11
Kluisbergen	0.21	0.20	0.08	0.24	0.17	0.23	0.14	0.17
Knesselare	0.33	0.30	0.21	0.02	0.09	0.15	0.15	0.18
Knokke-Heist	0.17	0.33	0.18	0.11	0.18	0.09	0.26	0.09
Koekelare	0.23	0.32	0.15	0.00	0.08	0.35	0.11	0.21
Koksijde	0.26	0.38	0.12	0.12	0.08	0.11	0.17	0.12
Kontich	0.27	0.26	0.21	0.12	0.14	0.17	0.21	0.11
Kortemark	0.17	0.23	0.15	0.15	0.08	0.23	0.11	0.23
Kortenaken	0.21	0.18	0.24	0.03	0.05	0.47	0.12	0.14
Kortenberg	0.26	0.18	0.11	0.21	0.11	0.21	0.18	0.17
Kortessem	0.27	0.30	0.12	0.23	0.02	0.23	0.12	0.11
Kortrijk	0.23	0.29	0.23	0.05	0.14	0.14	0.21	0.12
Kraainem	0.29	0.06	0.14	0.20	0.09	0.24	0.27	0.12
Kruibeke	0.26	0.18	0.29	0.09	0.06	0.23	0.18	0.12
Kruishoutem	0.29	0.21	0.08	0.21	0.15	0.30	0.14	0.09
Kuurne	0.17	0.23	0.14	0.23	0.17	0.15	0.18	0.18
Laakdal	0.27	0.18	0.12	0.29	0.03	0.27	0.15	0.15
Laarne	0.18	0.21	0.20	0.15	0.08	0.29	0.15	0.15
Lanaken	0.26	0.26	0.14	0.17	0.03	0.23	0.17	0.17
Landen	0.29	0.17	0.26	0.12	0.03	0.24	0.20	0.11
Langemark-Poelkapelle	0.33	0.26	0.15	0.08	0.08	0.29	0.09	0.15
Lebbeke	0.21	0.30	0.15	0.08	0.06	0.23	0.17	0.12
Lede	0.24	0.24	0.20	0.06	0.08	0.27	0.18	0.21
Ledegem	0.27	0.27	0.14	0.02	0.17	0.18	0.21	0.14
Lendelede	0.23	0.36	0.12	0.02	0.08	0.27	0.14	0.23

Lennik	0.18	0.12	0.09	0.18	0.08	0.35	0.30	0.12
Leopoldsburg	0.27	0.35	0.27	0.00	0.02	0.26	0.17	0.12
Leuven	0.26	0.27	0.26	0.09	0.11	0.11	0.26	0.08
Lichtervelde	0.21	0.24	0.14	0.21	0.03	0.20	0.14	0.23
Liedekerke	0.29	0.20	0.14	0.14	0.12	0.20	0.17	0.20
Lier	0.33	0.26	0.20	0.02	0.11	0.15	0.23	0.14
Lierde	0.29	0.36	0.09	0.02	0.08	0.23	0.14	0.21
Lille	0.23	0.12	0.08	0.20	0.12	0.17	0.14	0.29
Linkebeek	0.24	0.06	0.18	0.42	0.06	0.26	0.17	0.06
Lint	0.30	0.20	0.20	0.14	0.08	0.26	0.14	0.12
Linter	0.29	0.09	0.18	0.20	0.05	0.33	0.15	0.12
Lo-Reninge	0.44	0.26	0.06	0.00	0.11	0.32	0.12	0.12
Lochristi	0.18	0.17	0.15	0.27	0.09	0.27	0.12	0.21
Lokeren	0.24	0.18	0.17	0.24	0.11	0.12	0.20	0.20
Lommel	0.18	0.26	0.15	0.08	0.03	0.21	0.21	0.09
Londerzeel	0.14	0.18	0.20	0.11	0.08	0.29	0.27	0.12
Lovendegem	0.24	0.15	0.18	0.20	0.06	0.21	0.15	0.24
Lubbeek	0.27	0.17	0.21	0.29	0.02	0.17	0.14	0.18
Lummen	0.24	0.29	0.26	0.06	0.02	0.26	0.15	0.18
Maarkedal	0.33	0.14	0.17	0.24	0.09	0.21	0.15	0.15
Maaseik	0.20	0.35	0.23	0.06	0.03	0.27	0.14	0.09
Maasmechelen	0.23	0.23	0.35	0.08	0.02	0.23	0.18	0.12
Machelen	0.24	0.18	0.35	0.24	0.11	0.08	0.18	0.06
Maldegem	0.27	0.23	0.26	0.08	0.11	0.20	0.18	0.14
Malle	0.24	0.20	0.24	0.09	0.12	0.21	0.23	0.12
Mechelen	0.27	0.21	0.20	0.06	0.15	0.06	0.30	0.14
Meerhout	0.29	0.08	0.26	0.14	0.21	0.17	0.14	0.15
Meeuwen-Gruitrode	0.36	0.32	0.12	0.02	0.05	0.36	0.06	0.14
Meise	0.30	0.20	0.11	0.39	0.06	0.15	0.12	0.11
Melle	0.27	0.21	0.17	0.14	0.09	0.24	0.20	0.14
Menen	0.24	0.24	0.24	0.14	0.08	0.15	0.24	0.09
Merchtem	0.18	0.12	0.15	0.59	0.02	0.15	0.09	0.09
Merelbeke	0.21	0.24	0.20	0.11	0.11	0.23	0.15	0.18
Merksplas	0.18	0.35	0.17	0.17	0.12	0.24	0.09	0.11
Mesen	0.50	0.27	0.09	0.00	0.05	0.23	0.14	0.11
Meulebeke	0.24	0.29	0.18	0.00	0.15	0.27	0.17	0.14
Middelkerke	0.14	0.27	0.15	0.08	0.17	0.30	0.14	0.14
Moerbeke	0.20	0.15	0.23	0.24	0.06	0.20	0.12	0.20
Mol	0.41	0.18	0.24	0.09	0.09	0.08	0.20	0.15
Moorslede	0.21	0.39	0.09	0.15	0.05	0.26	0.11	0.18
Mortsel	0.27	0.17	0.17	0.39	0.11	0.08	0.15	0.06
Nazareth	0.20	0.32	0.23	0.18	0.08	0.21	0.14	0.12
Neerpelt	0.26	0.20	0.14	0.00	0.05	0.45	0.12	0.21
Nevele	0.23	0.17	0.30	0.18	0.12	0.20	0.18	0.09
Niel	0.44	0.14	0.12	0.23	0.03	0.26	0.15	0.11
Nieuwerkerken	0.35	0.15	0.17	0.05	0.02	0.27	0.15	0.15
Nieuwpoort	0.27	0.23	0.14	0.03	0.14	0.11	0.23	0.14
Nijlen	0.20	0.11	0.09	0.48	0.02	0.23	0.14	0.11
Ninove	0.15	0.15	0.18	0.24	0.09	0.30	0.18	0.20
Olen	0.32	0.14	0.18	0.15	0.06	0.20	0.14	0.24
Oostende	0.26	0.18	0.17	0.12	0.15	0.09	0.29	0.11
Oosterzele	0.24	0.15	0.15	0.15	0.06	0.26	0.12	0.21
Oostkamp	0.30	0.36	0.17	0.00	0.12	0.14	0.17	0.17

Oostrozebeke	0.36	0.24	0.24	0.00	0.15	0.20	0.15	0.12
Opglabbeek	0.45	0.29	0.14	0.00	0.05	0.20	0.12	0.15
Opwijk	0.20	0.15	0.17	0.45	0.05	0.14	0.09	0.12
Oud-Heverlee	0.21	0.17	0.15	0.48	0.02	0.20	0.12	0.11
Oud-Turnhout	0.29	0.15	0.14	0.12	0.09	0.18	0.20	0.23
Oudenaarde	0.17	0.29	0.17	0.02	0.12	0.18	0.23	0.20
Oudenburg	0.20	0.32	0.17	0.00	0.08	0.32	0.15	0.17
Overijse	0.29	0.18	0.20	0.24	0.06	0.15	0.18	0.11
Overpelt	0.33	0.23	0.32	0.02	0.18	0.23	0.08	0.09
Peer	0.42	0.18	0.23	0.00	0.05	0.33	0.11	0.09
Pepingen	0.26	0.21	0.14	0.00	0.05	0.36	0.17	0.26
Pittem	0.24	0.20	0.23	0.00	0.14	0.27	0.11	0.20
Poperinge	0.17	0.35	0.11	0.08	0.14	0.26	0.14	0.15
Putte	0.21	0.14	0.24	0.00	0.09	0.23	0.14	0.26
Puurs	0.24	0.20	0.17	0.00	0.18	0.29	0.18	0.15
Ranst	0.21	0.15	0.11	0.20	0.14	0.32	0.20	0.14
Ravels	0.26	0.18	0.11	0.27	0.03	0.32	0.12	0.20
Retie	0.24	0.12	0.18	0.21	0.05	0.23	0.12	0.27
Riemst	0.24	0.21	0.08	0.21	0.09	0.32	0.11	0.17
Rijkevorsel	0.29	0.11	0.14	0.08	0.05	0.50	0.14	0.14
Roeselare	0.21	0.24	0.26	0.18	0.12	0.17	0.20	0.05
Ronse	0.21	0.20	0.33	0.14	0.15	0.14	0.18	0.11
Roosdaal	0.29	0.14	0.14	0.26	0.08	0.21	0.18	0.15
Rotselaar	0.24	0.15	0.14	0.18	0.05	0.21	0.12	0.20
Ruiselede	0.32	0.30	0.11	0.03	0.05	0.27	0.15	0.23
Rumst	0.29	0.14	0.11	0.23	0.17	0.24	0.15	0.11
Schelle	0.26	0.12	0.20	0.27	0.12	0.24	0.14	0.11
Scherpenheuvel-Zichem	0.15	0.29	0.18	0.15	0.08	0.30	0.15	0.08
Schilde	0.27	0.17	0.14	0.06	0.09	0.29	0.20	0.21
Schoten	0.27	0.23	0.26	0.03	0.08	0.17	0.24	0.14
Sint-Amands	0.27	0.23	0.17	0.00	0.05	0.36	0.14	0.20
Sint-Genesius-Rode	0.21	0.09	0.24	0.20	0.11	0.21	0.23	0.18
Sint-Gillis-Waas	0.26	0.20	0.17	0.18	0.08	0.24	0.15	0.18
Sint-Katelijne-Waver	0.27	0.17	0.21	0.18	0.17	0.15	0.15	0.14
Sint-Laureins	0.21	0.18	0.11	0.20	0.06	0.26	0.17	0.29
Sint-Lievens-Houtem	0.27	0.23	0.12	0.08	0.05	0.15	0.20	0.21
Sint-Martens-Latem	0.20	0.24	0.11	0.29	0.06	0.23	0.18	0.15
Sint-Niklaas	0.27	0.21	0.32	0.14	0.06	0.06	0.21	0.17
Sint-Pieters-Leeuw	0.21	0.20	0.24	0.17	0.12	0.12	0.20	0.14
Sint-Truiden	0.23	0.29	0.21	0.08	0.11	0.14	0.18	0.11
Spiere-Helkijn	0.26	0.27	0.12	0.00	0.24	0.35	0.09	0.12
Stabroek	0.38	0.15	0.21	0.17	0.05	0.21	0.17	0.14
Staden	0.24	0.17	0.21	0.18	0.09	0.26	0.12	0.17
Steenokkerzeel	0.29	0.17	0.12	0.27	0.06	0.12	0.15	0.15
Stekene	0.29	0.18	0.18	0.12	0.15	0.11	0.14	0.21
Temse	0.33	0.21	0.21	0.03	0.12	0.12	0.17	0.24
Ternat	0.26	0.32	0.14	0.15	0.12	0.20	0.17	0.12
Tervuren	0.30	0.15	0.15	0.21	0.03	0.27	0.17	0.12
Tessenderlo	0.24	0.29	0.21	0.08	0.08	0.23	0.17	0.17
Tielt	0.24	0.36	0.23	0.03	0.06	0.21	0.12	0.14
Tielt-Winge	0.32	0.24	0.30	0.00	0.03	0.27	0.11	0.15
Tienen	0.20	0.29	0.20	0.06	0.08	0.20	0.23	0.14
Tongeren	0.48	0.24	0.06	0.15	0.03	0.18	0.14	0.09

Torhout	0.21	0.27	0.14	0.03	0.21	0.20	0.17	0.17
Tremelo	0.38	0.12	0.27	0.00	0.03	0.29	0.18	0.17
Turnhout	0.21	0.23	0.23	0.23	0.14	0.11	0.21	0.09
Veurne	0.29	0.24	0.20	0.12	0.11	0.20	0.14	0.14
Vilvoorde	0.32	0.14	0.20	0.29	0.12	0.06	0.27	0.08
Vleteren	0.30	0.33	0.05	0.00	0.05	0.36	0.14	0.18
Voeren	0.39	0.09	0.17	0.00	0.06	0.29	0.11	0.33
Vorselaar	0.36	0.14	0.18	0.17	0.11	0.20	0.14	0.14
Vosselaar	0.21	0.12	0.12	0.26	0.14	0.18	0.12	0.29
Waarschoot	0.30	0.26	0.15	0.02	0.09	0.30	0.15	0.09
Waasmunster	0.23	0.15	0.20	0.12	0.06	0.20	0.15	0.23
Wachtebeke	0.32	0.26	0.32	0.00	0.05	0.18	0.14	0.23
Waregem	0.20	0.29	0.20	0.24	0.06	0.14	0.12	0.20
Wellen	0.33	0.23	0.11	0.00	0.03	0.38	0.15	0.15
Wemmel	0.23	0.05	0.30	0.38	0.09	0.11	0.15	0.12
Wervik	0.20	0.33	0.20	0.00	0.20	0.26	0.15	0.06
Westerlo	0.24	0.14	0.18	0.21	0.08	0.24	0.17	0.21
Wetteren	0.27	0.26	0.18	0.12	0.12	0.14	0.17	0.14
Wevelgem	0.23	0.29	0.23	0.12	0.17	0.14	0.17	0.08
Wezembeek-Oppem	0.24	0.09	0.15	0.23	0.05	0.29	0.26	0.14
Wichelen	0.38	0.11	0.24	0.21	0.06	0.15	0.15	0.15
Wielsbeke	0.32	0.29	0.17	0.00	0.09	0.30	0.14	0.11
Wijnegem	0.23	0.23	0.23	0.32	0.09	0.12	0.17	0.05
Willebroek	0.35	0.18	0.29	0.00	0.09	0.20	0.21	0.12
Wingene	0.21	0.41	0.11	0.02	0.06	0.29	0.11	0.23
Wommelgem	0.29	0.21	0.15	0.21	0.11	0.23	0.18	0.12
Wortegem-Petegem	0.26	0.33	0.11	0.21	0.06	0.26	0.08	0.15
Wuustwezel	0.15	0.12	0.21	0.23	0.08	0.36	0.20	0.12
Zandhoven	0.23	0.18	0.08	0.26	0.09	0.24	0.17	0.11
Zaventem	0.29	0.14	0.17	0.18	0.15	0.09	0.35	0.08
Zedelgem	0.27	0.30	0.23	0.00	0.11	0.21	0.18	0.17
Zele	0.18	0.12	0.24	0.14	0.09	0.26	0.21	0.18
Zelzate	0.21	0.15	0.29	0.14	0.05	0.18	0.23	0.17
Zemst	0.26	0.12	0.12	0.24	0.11	0.20	0.14	0.24
Zingem	0.41	0.23	0.08	0.33	0.08	0.12	0.11	0.15
Zoersel	0.57	0.14	0.02	0.33	0.08	0.11	0.14	0.09
Zomergem	0.23	0.30	0.17	0.00	0.09	0.29	0.17	0.18
Zonhoven	0.35	0.27	0.24	0.00	0.02	0.27	0.18	0.12
Zonnebeke	0.24	0.38	0.12	0.15	0.15	0.23	0.12	0.08
Zottegem	0.54	0.17	0.17	0.20	0.05	0.05	0.20	0.12
Zoutleeuw	0.24	0.15	0.20	0.05	0.05	0.38	0.15	0.18
Zuienkerke	0.38	0.14	0.05	0.17	0.05	0.33	0.18	0.15
Zulte	0.29	0.23	0.09	0.11	0.12	0.33	0.17	0.12
Zutendaal	0.27	0.20	0.11	0.00	0.03	0.44	0.15	0.14
Zwalm	0.30	0.18	0.17	0.00	0.17	0.26	0.14	0.23
Zwevegem	0.23	0.27	0.20	0.23	0.09	0.23	0.11	0.08
Zwijndrecht	0.29	0.18	0.32	0.00	0.17	0.17	0.20	0.11

A.3 Spearman correlation tables

Table A.4: Spearman correlation among weight restriction models

Unconditional

	MinMax	Average	Municipal
MinMax restrictions	1.0000		
Average restrictions	0.8755	1.0000	
Municipal-specific restrictions	0.6796	0.7973	1.0000

Robust Unconditional

	MinMax	Average	Municipal
MinMax restrictions	1.0000		
Average restrictions	0.9202	1.0000	
Municipal-specific restrictions	0.8235	0.8913	1.0000

Robust Conditional Model 1

	MinMax	Average	Municipal
MinMax restrictions	1.0000		
Average restrictions	0.8834	1.0000	
Municipal-specific restrictions	0.8365	0.9074	1.0000

Robust Conditional Model 2

	MinMax	Average	Municipal
MinMax restrictions	1.0000		
Average restrictions	0.8052	1.0000	
Municipal-specific restrictions	0.7928	0.9423	1.0000

Robust Conditional Model 3

	MinMax	Average	Municipal
MinMax restrictions	1.0000		
Average restrictions	0.7629	1.0000	
Municipal-specific restrictions	0.7466	0.9279	1.0000

Note: MinMax restrictions refer to the minimum and maximum share of expenditure in each municipal area across all the municipalities. *Average* restrictions consider the average spending share (lower and upper bound equal to its $\pm 50\%$). *Municipal-specific* restrictions are based on the municipal-specific current spending allocation (lower and upper bound equal to the $\pm 50\%$ of each municipal spending share).

Unconditional refers to the baseline Benefit-of-the-Doubt Directional Distance function model, without correction for outlying observations, as it is instead in the *Robust* specification. *Model 1* includes the economic and financial characteristics (Fiscal income, Financial debt and Unemployment). *Model 2* adds to the economic and financial characteristics the socio-demographic structure (Share of elderly people, Share of foreigners and Population growth). *Model 3* adds to the economic-financial and socio-demographic variables also the political component (Ideological Complexion of the local Government). In every conditional model specification a year dummy is also included.

Table A.5: Spearman correlation among model specifications

MinMax weight restrictions

	Unconditional	Rob. Uncond.	Rob. Cond. 1	Rob. Cond. 2	Rob. Cond. 3
Unconditional	1.0000				
Rob. Uncond.	0.9553	1.0000			
Rob. Cond. 1	0.6708	0.7372	1.0000		
Rob. Cond. 2	0.2151	0.2475	0.3313	1.0000	
Rob. Cond. 3	0.1331	0.1592	0.2292	0.2930	1.0000

Average weight restrictions

	Unconditional	Rob. Uncond.	Rob. Cond. 1	Rob. Cond. 2	Rob. Cond. 3
Unconditional	1.0000				
Rob. Uncond.	0.9862	1.0000			
Rob. Cond. 1	0.7454	0.7643	1.0000		
Rob. Cond. 2	0.2118	0.2192	0.2914	1.0000	
Rob. Cond. 3	0.1129	0.1228	0.1978	0.3651	1.0000

Municipal-specific weight restrictions

	Unconditional	Rob. Uncond.	Rob. Cond. 1	Rob. Cond. 2	Rob. Cond. 3
Unconditional	1.0000				
Rob. Uncond.	0.9142	1.0000			
Rob. Cond. 1	0.6660	0.7548	1.0000		
Rob. Cond. 2	0.1682	0.2162	0.2986	1.0000	
Rob. Cond. 3	0.1040	0.1416	0.2152	0.3477	1.0000

Note: MinMax restrictions refer to the minimum and maximum share of expenditure in each municipal area across all the municipalities. *Average* restrictions consider the average spending share (lower and upper bound equal to its ±50%). *Municipal-specific* restrictions are based on the municipal-specific current spending allocation (lower and upper bound equal to the ±50% of each municipal spending share).

Unconditional refers to the baseline Benefit-of-the-Doubt Directional Distance function model, without correction for outlying observations, as it is instead in the *Robust* specification. *Model 1* includes the economic and financial characteristics (Fiscal income, Financial debt and Unemployment). *Model 2* adds to the economic and financial characteristics the socio-demographic structure (Share of elderly people, Share of foreigners and Population growth). *Model 3* adds to the economic-financial and socio-demographic variables also the political component (Ideological Complexion of the local Government). In every conditional model specification a year dummy is also included.

A.4 Shadow prices results

For the sake of brevity, only the shadow prices' descriptive statistics comparing the unrestricted unconditional and the robust conditional analysis for the three weight restriction models are presented below. Then, the budget shares obtained from the computed shadow prices are presented, again comparing the Unrestricted Unconditional and the robust conditional analysis for the three weight restriction models. For the sake of brevity, we show in the following only the estimates for the conditional model including the economic variables (Fiscal income, Financial debt and Unemployment). The results for the conditional model encompassing also Socio-demographic and Political components are available upon request from the authors.

For the sake of comparison, we recall below the average expenditure composition across the eight municipal functions considered in the present analysis.

	Administration	Culture	Care services	Education	Housing	Local mobility	Security	Environment
MIN	0.08	0.02	0.01	0	0.01	0.03	0.04	0.03
AVERAGE	0.18	0.14	0.12	0.10	0.05	0.15	0.11	0.10
MAX	0.49	0.31	0.23	0.39	0.16	0.33	0.23	0.22

We can observe that imposing weight restrictions gives back more reliable results, as closer with the current municipal composition. Budget allocation cannot be changed drastically, as otherwise suggested by the unrestricted results, and a minimum expenditure share is granted to each function in this way, avoiding zero weights whenever a minimum amount of investment is required in a certain municipal function.

For completeness, also the descriptive statistics related to the free variable v contained in the model introduced in Section 2.3 (formula 2.2) are presented.

Table A.6: Descriptive statistics of the shadow prices for each municipal function (Unrestricted unconditional and Robust conditional estimates)

Unrestricted unconditional estimates

	Administration	Culture	Care services	Education	Housing	Local mobility	Security	Environment	v
Mean	0.0906	0.0038	0.0667	0.0012	0.0004	0.0214	0.2042	0.9926	-0.1917
SD	0.0476	0.0042	0.1229	0.001	0.0006	0.0257	0.2882	0.3988	0.3611
Min	0	0	0	0	0	0	0	0	-0.7904
Max	0.2033	0.0166	0.6791	0.005	0.0028	0.1582	2.6924	1.8997	1.0795
N	307	307	307	307	307	307	307	307	307

MinMax weight restrictions

	Administration	Culture	Care services	Education	Housing	Local mobility	Security	Environment	v
Mean	0.4134	0.0093	0.1211	0.0016	0.0008	0.0277	0.2085	0.1832	0.4395
SD	0.1219	0.0069	0.0794	0.0015	0.0007	0.0183	0.1098	0.1013	0.1592
Min	0.1101	0.0025	0.0062	0	0.0001	0.0052	0.0549	0.0277	-0.0422
Max	0.6581	0.0348	0.3118	0.0065	0.0026	0.0994	0.5476	0.4386	0.8127
N	307	307	307	307	307	307	307	307	307

Average weight restrictions

	Administration	Culture	Care services	Education	Housing	Local mobility	Security	Environment	v
Mean	0.296	0.0146	0.1746	0.0016	0.0008	0.0335	0.2252	0.1975	0.3927
SD	0.0555	0.0041	0.0455	0.0005	0.0003	0.0118	0.0807	0.0567	0.1094
Min	0.1318	0.0079	0.0709	0.0005	0.0003	0.012	0.0628	0.0577	0.0295
Max	0.4483	0.0275	0.2675	0.0027	0.0016	0.0655	0.4308	0.3214	0.6256
N	307	307	307	307	307	307	307	307	307

Municipal-specific weight restrictions

	Administration	Culture	Care services	Education	Housing	Local mobility	Security	Environment	v
Mean	0.2948	0.0148	0.1709	0.0015	0.0008	0.0345	0.2229	0.2037	0.3816
SD	0.0992	0.006	0.0658	0.0012	0.0006	0.0186	0.0847	0.0899	0.1479
Min	0.0891	0.0029	0.0073	0	0.0001	0.0026	0.0704	0.0366	-0.0548
Max	0.7595	0.0592	0.3511	0.0067	0.0032	0.0981	0.6727	0.4956	0.7499
N	307	307	307	307	307	307	307	307	307

Table A.7: Descriptive statistics of the budget shares across the municipal functions (Unrestricted unconditional and Robust conditional estimates)

Unrestricted unconditional estimates

	Administration	Culture	Care services	Education	Housing	Local mobility	Security	Environment
Mean	7%	3%	6%	7%	3%	8%	9%	56%
SD	4%	3%	13%	7%	4%	9%	11%	21%
Min	0%	0%	0%	0%	0%	0%	0%	0%
Max	18%	19%	67%	47%	23%	51%	61%	89%
N	307	307	307	307	307	307	307	307

MinMax weight restrictions

	Administration	Culture	Care services	Education	Housing	Local mobility	Security	Environment
Mean	35%	6%	10%	10%	5%	12%	10%	11%
SD	9%	5%	6%	9%	5%	7%	5%	5%
Min	9%	2%	1%	0%	1%	3%	4%	3%
Max	49%	26%	23%	39%	16%	33%	23%	22%
N	307	307	307	307	307	307	307	307

Average weight restrictions

	Administration	Culture	Care services	Education	Housing	Local mobility	Security	Environment
Mean	25%	10%	14%	10%	5%	14%	11%	12%
SD	3%	3%	3%	3%	2%	4%	3%	2%
Min	10%	7%	6%	5%	3%	7%	5%	5%
Max	27%	21%	17%	14%	8%	22%	16%	15%
N	307	307	307	307	307	307	307	307

Municipal-specific weight restrictions

	Administration	Culture	Care services	Education	Housing	Local mobility	Security	Environment
Mean	25%	10%	14%	9%	5%	15%	11%	12%
SD	7%	4%	5%	8%	4%	7%	4%	5%
Min	9%	2%	1%	0%	1%	2%	4%	3%
Max	65%	34%	29%	42%	20%	34%	31%	29%
N	307	307	307	307	307	307	307	307

A.5 Statistical inference results

In the following, the statistical inference results are reported for each weight restriction specification:

1. *MinMax* restrictions refer to the minimum and maximum share of expenditure in each municipal area across all the municipalities;

2. *Average* restrictions consider the average spending share (lower and upper bound equal to its $\pm 50\%$);

3. *Municipal-specific* restrictions are based on the municipal-specific current spending allocation (lower and upper bound equal to the $\pm 50\%$ of each municipal spending share).

For the conditional model specification, three groups of background variables are considered (in every conditional model specification a year dummy is also included):

1. *Model 1* includes the economic and financial characteristics (Fiscal income, Financial debt and Unemployment);

2. *Model 2* adds to the economic and financial characteristics the socio-demographic structure (Share of elderly people, Share of foreigners and Population growth);

3. *Model 3* adds to the economic-financial and socio-demographic variables the political component (Ideological Complexion of the local Government).

For the sake of clarity, we say that the background variable has an unfavourable influence on the service provision assessment when the municipal composite indicator score increases only because the municipality under assessment is evaluated among similar municipalities: the service provision it can afford is lower compared with the one of municipalities facing a different context. The opposite holds when a background variable is found to have a favourable influence.

Table A.8: Influence of background conditions on municipal service composite indicator

MinMax weight restrictions

	Model 1			Model 2			Model 3		
	Influence	p-value		Influence	p-value		Influence	p-value	
Economic-financial									
Fiscal income	Unfavourable	0.000	***	Unfavourable	0.000	***	Unfavourable	0.000	***
Financial debt	Unfavourable	0.000	***	Favourable	0.000	***	Favourable	0.000	***
Unemployment	Unfavourable	0.000	***	Unfavourable	0.170		Unfavourable	0.085	*
Socio-demographic									
Residents over 65				Favourable	0.075	*	Favourable	0.000	***
Foreigners				Favourable	0.045	**	Favourable	0.080	*
Population growth				Unfavourable	0.000	***	Unfavourable	0.000	***
Political									
ICG							Unfavourable	0.000	***

Average weight restrictions

	Model 1			Model 2			Model 3		
	Influence	p-value		Influence	p-value		Influence	p-value	
Economic-financial									
Fiscal income	Unfavourable	0.000	***	Unfavourable	0.000	***	Unfavourable	0.000	***
Financial debt	Unfavourable	0.000	***	Unfavourable	0.000	***	Favourable	0.000	***
Unemployment	Unfavourable	0.000	***	Unfavourable	0.000	***	Unfavourable	0.000	***
Socio-demographic									
Residents over 65				Favourable	0.000	***	Favourable	0.210	
Foreigners				Favourable	0.100		Favourable	0.030	**
Population growth				Unfavourable	0.000	***	Unfavourable	0.000	***
Political									
ICG							Unfavourable	0.000	***

Municipal-specific weight restrictions

	Model 1			Model 2			Model 3		
	Influence	p-value		Influence	p-value		Influence	p-value	
Economic-financial									
Fiscal income	Unfavourable	0.000	***	Favourable	0.000	***	Favourable	0.000	***
Financial debt	Unfavourable	0.000	***	Unfavourable	0.000	***	Unfavourable	0.000	***
Unemployment	Unfavourable	0.000	***	Unfavourable	0.750		Unfavourable	0.005	***
Socio-demographic									
Residents over 65				Favourable	0.000	***	Favourable	0.000	***
Foreigners				Favourable	0.000	***	Favourable	0.035	**
Population growth				Unfavourable	0.000	***	Unfavourable	0.000	***
Political									
ICG							Unfavourable	0.000	***

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.6 Government size results

Table A.9 presents the optimal tax rate computed across the different model specifications and Figure A.1 shows the "inverted U-relationship" recurrent in every model specification.

Table A.9: Optimal tax rates across different model specifications

	Unconditional	Robust	Model 1	Model 2	Model 3
MinMax	2.04%	3.65%	5.30%	9.41%	9.51%
Average	5.00%	4.83%	5.29%	9.98%	9.36%
Municipal average	4.07%	3.58%	4.96%	12.15%	11.67%

Note: MinMax restrictions refer to the minimum and maximum share of expenditure in each municipal area across all the municipalities. *Average* restrictions consider the average spending share (lower and upper bound equal to its $\pm 50\%$). *Municipal-specific* restrictions are based on the municipal-specific current spending allocation (lower and upper bound equal to the $\pm 50\%$ of each municipal spending share).

Unconditional refers to the baseline Benefit-of-the-Doubt Directional Distance function model, without correction for outlying observations, as it is instead in the *Robust* specification. *Model 1* includes the economic and financial characteristics (Fiscal income, Financial debt and Unemployment). *Model 2* adds to the economic and financial characteristics the socio-demographic structure (Share of elderly people, Share of foreigners and Population growth). *Model 3* adds to the economic-financial and socio-demographic variables the political component (Ideological Complexion of the local Government). In every conditional model specification a year dummy is also included.

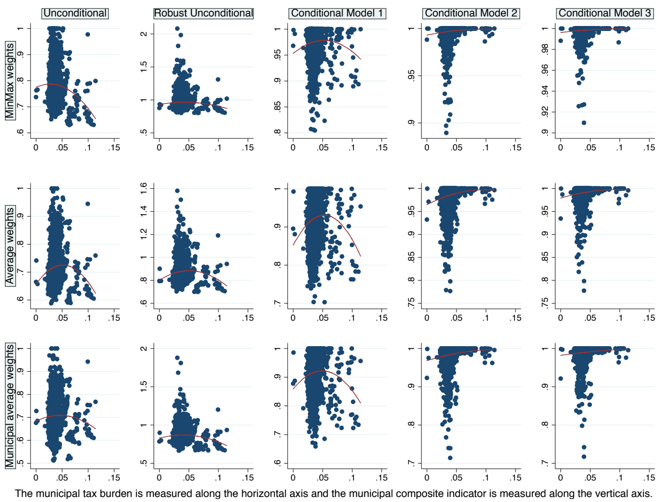


Figure A.1: "Inverted U-relationship" between municipal tax burden and municipal service provision level

A.7 Robustness check results

In the following, the main robustness check results are presented. Specifically, this further analysis has been performed excluding from the sample the 13 largest Flemish cities (so-called “Centrumsteden”), to make sure that the main findings are not influenced by the municipal size.

Table A.10: Descriptive statistics of the service provision composite indicator scores estimated for 294 municipalities over 2006–2011 (Analysis without “Centrumsteden”)

	Mean	St. Dev.	Min	Max
Unrestricted Unconditional	0.8324	0.0639	0.6261	1.0000
Unconditional				
MinMax restrictions	0.7704	0.0674	0.6219	1.0000
Average restrictions	0.7003	0.0632	0.5790	1.0000
Municipal-specific restrictions	0.6895	0.0713	0.4933	1.0000
Robust Unconditional				
MinMax restrictions	0.9501	0.1171	0.7768	2.1155
Average restrictions	0.8596	0.0973	0.6815	1.5953
Municipal-specific restrictions	0.8519	0.1104	0.6593	1.9019
Robust Conditional Model 1				
MinMax restrictions	0.9718	0.0355	0.7820	1.0009
Average restrictions	0.9142	0.0642	0.6893	1.0005
Municipal-specific restrictions	0.9084	0.0703	0.6558	1.0002
Robust Conditional Model 2				
MinMax restrictions	0.9967	0.0098	0.8622	1.0000
Average restrictions	0.9836	0.0283	0.7703	1.0000
Municipal-specific restrictions	0.9821	0.0319	0.7142	1.0000
Robust Conditional Model 3				
MinMax restrictions	0.9981	0.0063	0.9059	1.0000
Average restrictions	0.9900	0.0207	0.7771	1.0000
Municipal-specific restrictions	0.9887	0.0241	0.7194	1.0000

Note: Unrestricted indicates the absence of weight restrictions. *MinMax* restrictions refer to the minimum and maximum share of expenditure in each municipal area across all the municipalities. *Average* restrictions consider the average spending share (lower and upper bound equal to its $\pm 50\%$). *Municipal-specific* restrictions are based on the municipal-specific current spending allocation (lower and upper bound equal to the $\pm 50\%$ of each municipal spending share).

Model 1 includes the economic and financial characteristics (Fiscal income, Financial debt and Unemployment). *Model 2* adds to the economic and financial characteristics the socio-demographic structure (Share of elderly people, Share of foreigners and Population growth). *Model 3* adds to the economic-financial and socio-demographic variables the political component (Ideological Complexion of the local Government). In every conditional model specification a year dummy is also included.

Table A.11: Influence of background conditions on municipal service composite indicator (Analysis without "Centrumsteden")

MinMax weight restrictions

	Model 1			Model 2			Model 3		
	Influence	p-value		Influence	p-value		Influence	p-value	
Economic-financial									
Fiscal income	Unfavourable	0.000	***	Unfavourable	0.000	***	Unfavourable	0.000	***
Financial debt	Unfavourable	0.000	***	Favourable	0.000	***	Favourable	0.000	***
Unemployment	Unfavourable	0.000	***	Unfavourable	0.075	*	Unfavourable	0.155	
Socio-demographic									
Residents over 65				Favourable	0.995		Favourable	0.005	***
Foreigners				Favourable	0.885		Favourable	0.355	
Population growth				Unfavourable	0.000	***	Unfavourable	0.000	***
Political									
ICG							Unfavourable	0.000	***

Average weight restrictions

	Model 1			Model 2			Model 3		
	Influence	p-value		Influence	p-value		Influence	p-value	
Economic-financial									
Fiscal income	Unfavourable	0.000	***	Unfavourable	0.000	***	Unfavourable	0.000	***
Financial debt	Unfavourable	0.000	***	Favourable	0.000	***	Favourable	0.000	***
Unemployment	Unfavourable	0.000	***	Unfavourable	0.000	***	Unfavourable	0.000	***
Socio-demographic									
Residents over 65				Favourable	0.000	***	Favourable	0.000	***
Foreigners				Favourable	0.345		Favourable	0.570	
Population growth				Unfavourable	0.005	***	Unfavourable	0.000	***
Political									
ICG							Unfavourable	0.000	***

Municipal-specific weight restrictions

	Model 1			Model 2			Model 3		
	Influence	p-value		Influence	p-value		Influence	p-value	
Economic-financial									
Fiscal income	Unfavourable	0.000	***	Favourable	0.000	***	Favourable	0.000	***
Financial debt	Unfavourable	0.000	***	Unfavourable	0.000	***	Unfavourable	0.000	***
Unemployment	Unfavourable	0.000	***	Unfavourable	0.010	**	Unfavourable	0.000	***
Socio-demographic									
Residents over 65				Favourable	0.025	**	Favourable	0.000	***
Foreigners				Favourable	0.005	***	Favourable	0.005	***
Population growth				Unfavourable	0.000	***	Unfavourable	0.000	***
Political									
ICG							Unfavourable	0.000	***

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

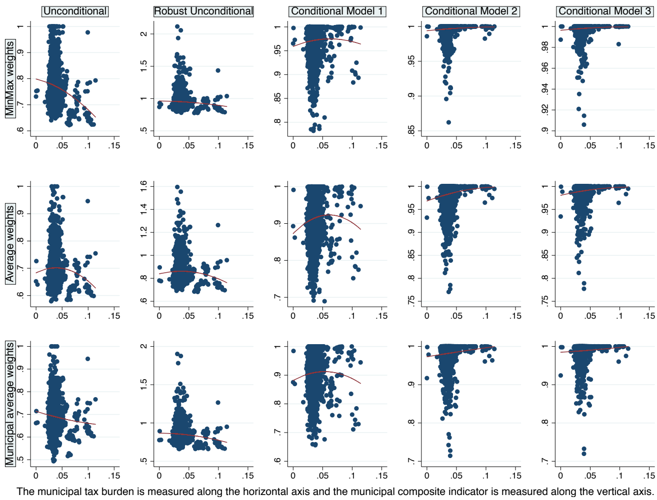


Figure A.2: "Inverted U-relationship" between municipal tax burden and municipal service provision level (Analysis without "Centrumsteden")

Table A.10 shows the descriptive statistics of the service provision composite indicator scores estimated for 294 municipalities over 2006–2011. Table A.11 presents the statistical influence results for each weight restriction and each conditional model specification. Figure A.2 shows the "inverted U-relationship" recurrent in every model specification. The presented results confirm the findings obtained performing the analysis over the whole sample.

Appendix B

Supplementary material for Chapter 3

B.1 Some essential notions of wastewater treatment

To understand the background of the present analysis, a short description of wastewater treatment is provided. Due to the breadth of the subject, only the information strictly related to those technical features, explicitly recalled in the previous sections is provided. For a comprehensive and detailed presentation the reader can refer for example to Metcalf & Eddy et al. (2003).
Wastewater may be defined as a mix of liquid or water-borne wastes discharged by houses, commercial properties, factories, farms and public institutions. Ground, surface and stormwaters may be also included as components of the wastewater flow. As a preliminary step, wastewater is collected in sewers and then conveyed to treatment or disposal facilities. Regarding sewerage, two different systems can be identified: combined or separate sewerage. The first one transports both stormwater and wastewater, while separate sewerage is designed either to convey wastewater (sanitary sewers), or to drain surface runoff (storm sewers). For both combined and separate sewerage, the estimated dry weather flow is a relevant value. In fact, during the period of dry weather, infiltration and surface runoff have a minimum influence in combined sew-

erage. Therefore the estimated dry weather flow provides a basis for works design. Wastewater entering in the treatment plant is called influent or ingoing water and it is characterized by the presence of several physical, chemical and biological constituents such as suspended solids (Ss), nitrogen (N), phosphorus (P), carbon (C) and biodegradable organics; the latter is usually measured in terms of biochemical oxygen demand (BOD) and of chemical oxygen demand (COD). Together with constituents, pathogens and priority pollutants (e.g. heavy metals, pharmaceutical molecules) can be found in the influent wastewater. One of the main objectives of a wastewater treatment plant (WWTP) should be the removal of as many contaminants as possible and the purification of the water for further reuse. A suitable wastewater treatment should guarantee an acceptable level of overall water quality and this can be done by using different methods. The treatment methods which exploit physical phenomena are referred as Unit Operations (UO); on the other hand, if the removal of contaminants is based on chemical reactions, then the process is called Unit Process (UP).
To ensure a certain level of contaminant removal, unit operations and processes are jointly performed according to different wastewater treatment technologies:

- preliminary treatment removes gross solids (large objects, rags, grit);
- primary treatment eliminates floating and settable materials;
- advanced primary treatment allows the removal of suspended solids;
- secondary treatment eliminates organic contaminants thanks to biological and chemical processes which are carried out;
- tertiary treatment eliminates other pollutants that cannot be removed by means of primary and secondary treatments.

The secondary treatment removes also nitrogen and phosphorus (in this context they are also referred as nutrients) together with other pathogens and some heavy metals (Metcalf & Eddy et al., 2003). Among the different technologies that can be used in the secondary treatment, it is worth mentioning the activated sludge process. This can be considered as the main biological process in secondary treatment and it basically refers to a mass of microorganisms metabolizing the suspended and soluble matter in an aeration basin. Solids and in particular biosolids can be considered by-products of the wastewater treatment; they are often referred to as “sludge” and they are by far the largest removed constituents. After a primary treatment, solids can be further biologically, chemically or by heat treated (e.g. stabilization, composting, dewatering, drying, thickening) so to get them suitable for reuse (agriculture, home gardens...). The term biosolids indicates that solids are further treated and an important distinction between class A biosolids and class B biosolids has to be made (Metcalf & Eddy et al., 2003). Biosolids belonging to the first class are also known as "clean sludge" while Class B biosolids have a reduced concentration of pathogens and other unhealthy contaminants (mainly metals), but they do not satisfy specific legal requirements and therefore their application to land is strictly regulated. Sludge which is not eligible for further use, is then transported to either landfill or incinerators. Moreover, among the other constituents, nitrogen occupies a preeminent position in wastewater treatment activities; excessive concentrations of nitrogen can be harmful to humans and wildlife. Nitrogen can be found in the wastewater under various forms, namely organic nitrogen and inorganic nitrogen which is in turn divided into ammonia nitrogen, nitrite nitrogen and nitrate nitrogen. Ammonia concentrations can affect hatching and growth rates of fish. If excessive amounts of nitrates are discharged into the aquatic environment, it can lead to the growth of undesirable aquatic life and then to eutrophication. Nitrate can even affect human health if it is present in drinking water. Moreover, a great discharged of total nitrogen onto land can lead to the pollution of groundwater, causing excessive vegetative growth and a reduction of crop quality. Due to this, many alternative technologies have been designed to remove total nitrogen from wastewater (suspended growth nitrification and denitrification variations, attached growth nitrification and denitrification variations, biological nutrient removal variations). On the other hand, it is almost impossible to completely remove nitrogen from wastewater remaining in the effluent flow.

B.2 Water Performance Indexes results

In the following, the Water efficiency Performance Index (WPI) results for 96 wastewater treatment plants are reported, according to different model specifications and two sets of weights: the label "Non-radial" refers to the set of weights proposed by Zhou et al. (2012), while the label "AHP-non-radial" is related to the set of weights obtained applying the Analytic Hierarchy Process (AHP) approach. The indexes WPI_1, WPI_2 and WPI_3 have been defined according to the following model specifications:

Table B.1: Indexes

	Input model	Input/Good Out. model	Input/Good/Bad Out. model
β	$\beta=(\beta_x)$	$\beta=(\beta_x,\beta_y)$	$\beta=(\beta_x,\beta_y,\beta_b)$
g	$g=(-x,0,0)$	$g=(-x,y,0)$	$g=(-x,y,-b)$
Index	$WPI_1 = 1-\sum_{n=1}^{N} w_{x_n}\beta_{x_n}$	$WPI_2 = \dfrac{1-\sum_{n=1}^{N} w_{x_n}\beta_{x_n}}{1+\sum_{m=1}^{M} w_{y_m}\beta_{y_m}}$	$WPI_3 = \dfrac{1-\left(\sum_{n=1}^{N} w_{x_n}\beta_{x_n}+\sum_{j=1}^{J} w_{b_j}\beta_{b_j}\right)}{1+\sum_{m=1}^{M} w_{y_m}\beta_{y_m}}$

WPI_0 has been constructed by considering a non-radial Directional Distance Function model where the undesirable output is completely ignored for the sake of comparison with the other model specifications. Referring to the parameter specification, the form of β and g has been set as $\beta = (\beta_x, \beta_y)$ and $g = (-x, y)$, while the equality constraint associated with the weak disposability of the undesirable output has been cancelled. The set of weights of WPI_0 coincides with the one for WPI_2.

Table B.2: Water efficiency Performance Indexes results

	Non-radial				AHP-non-radial			
WWTP	WPI_0	WPI_1	WPI_2	WPI_3	WPI_0	WPI_1	WPI_2	WPI_3
1	0.726	0.776	0.735	0.778	0.652	0.695	0.688	0.791
2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.431	1.000	1.000	1.000	0.423	1.000	1.000	1.000
6	0.282	1.000	1.000	1.000	0.347	1.000	1.000	1.000
7	0.342	0.545	0.367	0.434	0.397	0.463	0.437	0.574
8	0.662	0.792	0.790	0.893	0.772	0.903	0.903	0.959
9	0.190	0.402	0.232	0.260	0.223	0.342	0.297	0.565
10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
11	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
12	0.082	0.473	0.083	0.090	0.236	0.441	0.237	0.411
13	0.286	0.241	0.252	0.540	0.208	0.241	0.242	0.686
14	0.198	0.184	0.197	0.332	0.154	0.208	0.220	0.474
15	0.291	0.404	0.409	0.683	0.251	0.492	0.492	0.814
16	0.426	0.430	0.440	0.436	0.350	0.480	0.480	0.509
17	0.709	1.000	1.000	1.000	0.706	1.000	1.000	1.000
18	0.287	1.000	1.000	1.000	0.184	1.000	1.000	1.000
19	0.415	0.417	0.423	0.400	0.290	0.418	0.335	0.381
20	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
21	0.192	0.415	0.356	0.439	0.161	0.423	0.423	0.695
22	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
23	0.236	0.240	0.272	0.325	0.151	0.248	0.241	0.342
24	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
25	0.647	1.000	1.000	1.000	0.561	1.000	1.000	1.000
26	0.355	0.465	0.365	0.284	0.324	0.445	0.358	0.376
27	0.397	0.416	0.473	0.575	0.265	0.428	0.428	0.632
28	0.247	0.501	0.253	0.268	0.419	0.518	0.497	0.454
29	0.631	0.697	0.697	0.728	0.551	0.705	0.705	0.666
30	0.321	0.522	0.343	0.395	0.414	0.457	0.451	0.605
31	0.195	0.395	0.220	0.267	0.231	0.307	0.308	0.589
32	0.366	0.431	0.417	0.483	0.225	0.415	0.385	0.585
33	0.403	0.550	0.594	0.653	0.271	0.580	0.580	0.709
34	0.194	0.400	0.241	0.312	0.219	0.284	0.287	0.548
35	0.497	1.000	1.000	1.000	0.522	1.000	1.000	1.000
36	0.469	0.525	0.573	0.539	0.366	0.546	0.484	0.515
37	0.186	0.282	0.307	0.324	0.107	0.287	0.219	0.357
38	0.374	1.000	1.000	1.000	0.677	1.000	1.000	1.000
39	0.252	0.589	0.333	0.370	0.437	0.630	0.630	0.774
40	0.399	1.000	1.000	0.304	0.335	1.000	1.000	0.245
41	0.714	1.000	1.000	0.580	0.822	1.000	1.000	0.446
42	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
43	0.048	0.158	0.048	0.050	0.099	0.173	0.134	0.122
44	0.375	0.521	0.456	0.574	0.543	0.576	0.547	0.773
45	0.338	1.000	1.000	1.000	0.515	1.000	1.000	1.000
46	0.182	0.462	0.311	0.444	0.138	0.554	0.286	0.569
47	0.179	1.000	1.000	1.000	0.125	1.000	1.000	1.000
48	0.027	0.585	0.048	0.049	0.085	0.527	0.141	0.174

49	0.293	0.399	0.431	0.483	0.201	0.438	0.380	0.521
50	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
51	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
52	0.279	1.000	1.000	1.000	0.332	1.000	1.000	1.000
53	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
54	0.335	0.477	0.354	0.341	0.243	0.542	0.317	0.289
55	0.144	0.410	0.222	0.322	0.123	0.267	0.234	0.565
56	0.150	0.246	0.167	0.183	0.113	0.229	0.227	0.493
57	0.067	0.261	0.074	0.097	0.098	0.187	0.130	0.179
58	0.052	0.158	0.056	0.054	0.070	0.149	0.124	0.215
59	0.126	0.242	0.154	0.214	0.088	0.225	0.212	0.534
60	0.286	0.509	0.476	0.555	0.163	0.367	0.331	0.603
61	0.438	0.479	0.459	0.431	0.292	0.516	0.376	0.370
62	0.262	0.634	0.344	0.194	0.402	0.780	0.407	0.240
63	0.415	0.669	0.427	0.395	0.405	0.701	0.480	0.345
64	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
65	0.595	0.649	0.605	0.505	0.521	0.693	0.539	0.412
66	0.275	0.587	0.332	0.320	0.312	0.565	0.543	0.562
67	0.335	1.000	1.000	1.000	0.464	1.000	1.000	1.000
68	0.079	0.429	0.112	0.104	0.086	0.346	0.175	0.171
69	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
70	0.216	0.525	0.227	0.174	0.298	0.583	0.398	0.318
71	0.399	0.679	0.405	0.341	0.267	0.737	0.310	0.242
72	0.136	0.506	0.200	0.219	0.260	0.508	0.418	0.690
73	0.397	1.000	1.000	1.000	0.365	1.000	1.000	1.000
74	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
75	0.157	0.395	0.280	0.350	0.087	0.338	0.270	0.500
76	0.228	0.500	0.274	0.233	0.170	0.503	0.453	0.487
77	0.317	0.534	0.328	0.289	0.233	0.619	0.314	0.268
78	0.159	0.522	0.310	0.375	0.085	0.348	0.185	0.307
79	0.233	0.553	0.364	0.408	0.159	0.552	0.365	0.403
80	0.312	0.453	0.389	0.365	0.206	0.368	0.308	0.349
81	0.107	0.400	0.165	0.136	0.082	0.312	0.157	0.217
82	0.049	0.365	0.050	0.055	0.114	0.371	0.130	0.067
83	0.287	0.486	0.382	0.380	0.171	0.427	0.298	0.343
84	0.259	0.507	0.259	0.255	0.166	0.497	0.198	0.187
85	0.353	0.541	0.510	0.580	0.262	0.611	0.477	0.490
86	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
87	0.158	0.329	0.204	0.237	0.089	0.304	0.164	0.218
88	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
89	0.138	0.434	0.173	0.129	0.090	0.438	0.219	0.242
90	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
91	0.095	0.504	0.235	0.277	0.049	0.413	0.185	0.273
92	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
93	0.058	0.615	0.212	0.234	0.032	0.548	0.160	0.236
94	0.092	0.622	0.266	0.239	0.063	0.500	0.304	0.456
95	0.131	0.604	0.312	0.374	0.069	0.520	0.184	0.309
96	0.297	0.741	0.431	0.457	0.235	0.797	0.457	0.370

Appendix C

Supplementary material for Chapter 4

C.1 Municipal efficiency scores and composite indicators at local and global level

Table C.1: Municipal efficiency scores and composite indicators at local and global level

Municipality	General administation	Local police	Educational services	Road maintenance	Social services	DEA-like overall	Municipal weights	Tuscan weights
Abbadia San Salvatore	0.683	0.422	0.421	0.263	0.772	0.490	0.561	0.589
Abetone	0.215	0.020	0.316	0.032	0.394	0.197	0.139	0.236
Agliana	0.969	0.604	0.504	0.430	0.427	0.567	0.622	0.671
Altopascio	0.681	0.522	0.770	0.842	0.952	0.751	0.742	0.766
Anghiari	0.654	0.453	0.508	0.224	0.177	0.406	0.409	0.451
Arcidosso	0.474	0.352	0.297	0.294	0.511	0.371	0.416	0.422
Arezzo	0.845	1.000	0.785	1.000	1.000	0.918	0.903	0.902
Asciano	0.556	0.572	0.369	0.304	0.441	0.439	0.470	0.469
Aulla	0.599	0.750	0.310	0.209	0.873	0.521	0.511	0.580
Badia Tedalda	0.634	0.231	0.181	0.138	0.050	0.229	0.246	0.331
Bagni di Lucca	0.458	0.343	0.305	0.170	0.578	0.356	0.383	0.417
Bagno a Ripoli	0.676	0.467	0.359	0.371	0.695	0.489	0.569	0.574
Bagnone	0.267	0.241	0.346	0.133	0.467	0.291	0.269	0.309
Barberino di Mugello	0.605	0.538	0.398	0.336	0.414	0.449	0.493	0.489
Barberino Val d'Elsa	0.328	0.760	0.510	0.543	0.196	0.486	0.354	0.388
Barga	0.514	0.380	0.225	0.246	1.000	0.438	0.439	0.536
Bibbiena	0.760	0.828	0.506	0.336	0.687	0.608	0.649	0.655
Bibbona	0.294	0.419	0.190	0.331	0.231	0.288	0.279	0.277

Bientina	0.774	0.511	0.497	0.478	0.242	0.495	0.537	0.547
Borgo a Mozzano	0.517	0.440	0.423	0.187	0.517	0.410	0.438	0.455
Borgo San Lorenzo	0.753	0.606	0.492	0.241	0.460	0.500	0.543	0.568
Buggiano	0.706	0.670	0.380	0.738	0.374	0.559	0.556	0.575
Buonconvento	0.468	0.253	0.405	0.369	0.135	0.330	0.344	0.351
Buti	0.693	0.356	0.537	0.278	0.287	0.429	0.500	0.495
Calci	0.557	0.473	0.477	0.345	0.239	0.421	0.451	0.437
Calcinaia	0.693	0.531	0.651	0.452	0.457	0.560	0.602	0.589
Calenzano	0.652	0.404	0.526	0.460	0.444	0.493	0.537	0.540
Camaiore	0.659	0.373	0.526	0.288	0.571	0.475	0.545	0.549
Campagnatico	0.488	0.258	0.475	0.170	0.310	0.344	0.395	0.388
Campi Bisenzio	0.563	0.683	1.000	0.362	1.000	0.741	0.693	0.723
Campiglia Marittima	0.643	0.381	0.565	0.573	0.505	0.530	0.565	0.569
Campo nell'Elba	0.249	0.374	0.271	0.172	0.598	0.325	0.264	0.335
Camporgiano	0.895	0.159	0.306	0.159	0.602	0.387	0.471	0.581
Cantagallo	0.407	0.407	0.945	0.199	0.242	0.484	0.387	0.433
Capalbio	0.436	0.323	0.268	0.364	0.284	0.326	0.367	0.355
Capannoli	0.757	0.668	0.661	0.636	0.321	0.614	0.626	0.618
Capannori	0.526	1.000	0.616	1.000	0.707	0.773	0.630	0.678
Capoliveri	0.301	0.257	0.298	0.215	0.384	0.288	0.290	0.306
Capolona	0.846	0.653	0.380	0.314	0.399	0.497	0.571	0.585
Capraia e Limite	0.811	1.000	0.599	0.477	0.391	0.654	0.659	0.653
Capraia Isola	0.488	0.071	0.447	1.000	1.000	0.581	0.467	0.629
Caprese Michelangelo	0.622	0.158	0.258	0.141	0.241	0.266	0.380	0.377
Careggine	0.470	0.251	0.748	0.181	0.503	0.448	0.440	0.471
Carmignano	0.945	0.556	0.496	0.517	0.437	0.571	0.638	0.669
Carrara	0.735	0.585	0.607	0.363	0.956	0.633	0.673	0.708
Casale Marittimo	0.552	0.135	0.393	1.000	0.159	0.445	0.428	0.455
Casciana Terme	0.526	0.419	0.616	0.393	0.492	0.496	0.516	0.508
Cascina	0.807	0.952	0.574	1.000	0.881	0.824	0.793	0.820
Casola in Lunigiana	0.529	0.067	0.278	0.101	0.416	0.261	0.346	0.373
Casole d'Elsa	0.473	0.551	0.259	0.298	0.203	0.349	0.357	0.359
Castagneto Carducci	0.449	0.319	0.344	0.486	0.295	0.375	0.384	0.390
Castel del Piano	0.514	0.546	0.411	0.274	0.711	0.478	0.488	0.516
Castel Focognano	0.664	0.465	0.286	0.150	0.361	0.366	0.419	0.453
Castel San Niccol	0.630	0.579	0.284	0.181	0.112	0.347	0.329	0.394
Castelfiorentino	0.838	0.612	0.404	0.621	0.831	0.629	0.704	0.720
Castelfranco di Sopra	0.720	0.415	0.431	0.237	0.070	0.371	0.346	0.438
Castelfranco di Sotto	0.596	0.689	1.000	0.405	0.533	0.676	0.601	0.633
Castell'Azzara	0.414	0.220	0.238	0.185	0.378	0.275	0.345	0.333
Castellina in Chianti	0.445	0.261	0.296	0.422	0.247	0.328	0.363	0.357
Castellina Marittima	0.499	0.305	0.361	0.295	0.255	0.339	0.411	0.379
Castelnuovo Berardenga	0.619	0.862	0.389	0.569	0.194	0.524	0.451	0.495
Castelnuovo di Garfagnana	0.526	0.289	0.366	0.316	0.210	0.338	0.385	0.381
Castelnuovo di Val di Cecina	0.435	0.329	0.335	0.711	0.227	0.406	0.388	0.395
Castiglion Fibocchi	0.541	0.192	0.585	0.140	0.458	0.387	0.432	0.453
Castiglione d'Orcia	0.442	0.266	0.316	0.272	0.514	0.350	0.388	0.403
Castiglione della Pescaia	0.205	0.192	0.254	0.143	0.261	0.213	0.204	0.217
Castiglione di Garfagnana	0.700	0.451	0.364	0.149	0.408	0.398	0.459	0.490
Cavriglia	0.553	0.733	0.492	0.532	0.294	0.525	0.490	0.494
Cecina	0.696	0.501	0.581	0.319	0.702	0.550	0.620	0.617
Cerreto Guidi	0.872	0.508	0.370	0.611	0.398	0.527	0.574	0.618
Certaldo	0.705	0.643	0.366	0.264	0.816	0.531	0.590	0.616

Cetona	0.502	0.287	0.379	0.343	0.318	0.361	0.420	0.403
Chianciano Terme	0.475	0.302	0.270	0.269	0.671	0.376	0.410	0.448
Chianni	0.455	0.171	0.196	0.160	0.301	0.241	0.313	0.318
Chiesina Uzzanese	0.816	0.281	0.544	0.323	0.521	0.483	0.594	0.600
Chitignano	0.792	0.073	0.498	0.258	0.161	0.350	0.486	0.474
Chiusdino	0.417	1.000	0.243	0.267	0.290	0.438	0.350	0.387
Chiusi	0.678	0.370	0.473	0.275	0.739	0.489	0.571	0.585
Chiusi della Verna	0.496	0.243	0.170	0.149	0.392	0.268	0.325	0.355
Cinigiano	0.493	0.203	0.215	0.178	0.761	0.341	0.376	0.448
Civitella in Val di Chiana	0.868	1.000	0.465	0.399	0.169	0.572	0.467	0.592
Civitella Paganico	0.330	0.205	0.250	0.201	0.111	0.219	0.239	0.240
Colle di Val d’Elsa	0.893	0.555	0.680	0.299	1.000	0.666	0.759	0.784
Collesalvetti	0.661	0.576	0.399	0.403	0.683	0.525	0.579	0.585
Comano	0.699	0.143	0.330	0.179	0.198	0.295	0.423	0.414
Coreglia Antelminelli	0.542	0.758	0.528	0.233	0.254	0.471	0.448	0.452
Cortona	0.795	0.688	0.340	0.283	0.715	0.533	0.607	0.630
Crespina	0.527	0.353	0.333	0.267	0.273	0.343	0.406	0.390
Cutigliano	0.327	0.163	0.268	0.079	0.727	0.296	0.253	0.367
Dicomano	0.860	0.789	0.335	0.250	0.318	0.488	0.514	0.567
Empoli	1.000	0.552	0.506	0.617	0.683	0.643	0.731	0.762
Fabbriche di Vallico	0.345	0.068	0.854	0.180	0.158	0.362	0.322	0.344
Fauglia	0.494	0.525	0.400	0.410	0.530	0.464	0.471	0.479
Fiesole	0.420	0.328	0.505	0.317	0.830	0.474	0.473	0.510
Figline Valdarno	0.784	0.295	0.669	0.334	0.540	0.521	0.584	0.615
Filattiera	0.552	0.157	0.289	0.140	0.367	0.286	0.373	0.384
Firenzuola	0.566	0.646	0.251	0.277	0.334	0.399	0.408	0.431
Fivizzano	0.616	0.285	0.207	0.210	0.841	0.394	0.445	0.525
Foiano della Chiana	0.753	0.514	0.289	0.584	0.340	0.473	0.489	0.540
Follonica	0.402	0.368	0.630	0.341	0.827	0.518	0.487	0.529
Forte dei Marmi	0.162	0.104	0.131	0.167	0.205	0.150	0.156	0.163
Fosciandora	0.399	1.000	0.754	0.348	0.173	0.570	0.410	0.447
Fosdinovo	0.641	0.582	0.439	0.315	1.000	0.570	0.571	0.647
Fucecchio	0.865	0.599	0.652	0.363	0.833	0.647	0.744	0.740
Gaiole in Chianti	0.459	0.332	0.229	0.491	0.408	0.368	0.384	0.402
Gallicano	0.371	0.265	0.283	0.172	0.181	0.253	0.292	0.280
Gambassi Terme	0.825	0.511	0.337	0.422	0.324	0.462	0.546	0.554
Gavorrano	0.620	0.660	0.420	0.332	0.850	0.556	0.574	0.609
Giuncugnano	0.782	0.077	0.493	0.232	0.279	0.363	0.527	0.494
Greve in Chianti	0.749	1.000	0.401	1.000	0.563	0.724	0.664	0.698
Grosseto	0.938	0.575	0.525	0.994	0.782	0.736	0.773	0.811
Guardistallo	0.480	0.162	0.438	0.274	0.284	0.329	0.411	0.377
Impruneta	0.568	0.542	0.531	0.319	0.701	0.526	0.561	0.560
Incisa in Val d’Arno	0.703	0.413	0.330	0.313	0.239	0.385	0.444	0.463
Isola del Giglio	0.226	0.049	0.603	0.296	0.818	0.408	0.248	0.421
Lajatico	0.458	0.190	0.468	0.202	0.314	0.330	0.382	0.374
Lamporecchio	1.000	0.618	0.328	0.398	0.589	0.548	0.633	0.689
Larciano	0.719	0.478	0.408	0.388	0.452	0.473	0.555	0.546
Lari	0.556	0.357	0.317	0.402	0.532	0.415	0.463	0.476
Lastra a Signa	0.581	0.763	0.552	0.893	0.595	0.675	0.602	0.632
Laterina	0.673	0.453	0.453	0.325	0.434	0.458	0.541	0.521
Licciana Nardi	0.623	0.300	0.264	0.213	0.396	0.338	0.421	0.435
Livorno	0.697	0.935	0.665	0.674	0.596	0.715	0.677	0.684
Londa	0.521	0.384	0.634	0.112	0.181	0.384	0.388	0.400

Lorenzana	0.662	0.104	0.425	0.276	0.341	0.351	0.481	0.457
Loro Ciuffenna	0.771	0.464	0.365	0.254	0.251	0.406	0.473	0.496
Lucca	0.935	0.613	0.744	0.559	0.472	0.663	0.661	0.724
Lucignano	0.549	0.278	0.409	0.277	0.101	0.324	0.331	0.367
Magliano in Toscana	0.425	0.249	0.377	0.487	0.455	0.393	0.408	0.418
Manciano	0.498	0.467	0.382	0.277	1.000	0.502	0.493	0.567
Marciana	0.256	0.167	0.222	0.138	0.193	0.194	0.218	0.214
Marciana Marina	0.227	0.080	0.321	0.151	0.244	0.210	0.216	0.226
Marciano della Chiana	0.812	0.397	0.373	0.328	0.527	0.462	0.563	0.581
Marliana	0.557	0.428	0.361	0.399	0.320	0.405	0.462	0.440
Marradi	0.584	0.290	0.485	0.187	0.203	0.352	0.401	0.407
Massa	0.624	1.000	0.742	0.925	0.693	0.805	0.703	0.726
Massa e Cozzile	0.867	0.479	0.408	0.437	0.523	0.518	0.618	0.628
Massa Marittima	0.553	0.383	0.558	0.369	0.533	0.479	0.515	0.513
Massarosa	0.582	0.588	0.425	0.611	0.474	0.528	0.527	0.535
Minucciano	0.478	0.690	0.253	0.117	0.351	0.366	0.344	0.384
Molazzana	0.518	0.123	0.483	0.142	0.184	0.295	0.365	0.358
Monsummano Terme	0.780	0.503	0.505	0.574	0.575	0.572	0.637	0.640
Montaione	0.397	0.344	0.290	0.606	0.020	0.334	0.127	0.313
Montalcino	0.306	0.244	0.370	0.168	0.147	0.256	0.247	0.258
Montale	0.763	0.430	0.307	0.713	0.400	0.498	0.530	0.570
Monte Argentario	0.379	0.205	0.597	0.505	0.517	0.451	0.403	0.448
Monte San Savino	0.792	0.442	0.567	0.280	0.516	0.510	0.593	0.601
Montecarlo	0.899	0.509	0.656	0.334	0.293	0.536	0.603	0.618
Montecatini Val di Cecina	0.424	0.577	0.237	0.927	0.388	0.498	0.450	0.457
Montecatini-Terme	0.449	0.206	0.428	0.202	0.963	0.432	0.450	0.516
Montelupo Fiorentino	0.830	0.534	0.396	0.264	0.605	0.500	0.577	0.614
Montemignaio	0.502	0.143	0.342	0.114	0.074	0.234	0.293	0.300
Montemurlo	0.512	0.409	0.497	0.334	0.596	0.465	0.503	0.499
Montepulciano	0.566	0.505	0.464	0.284	0.570	0.470	0.513	0.511
Monterchi	0.573	0.213	0.329	0.173	0.482	0.337	0.413	0.434
Monteriggioni	0.675	0.527	0.337	0.298	0.359	0.423	0.477	0.488
Monteroni d'Arbia	0.696	0.384	0.294	0.485	0.192	0.394	0.384	0.462
Montescudaio	0.480	0.062	0.590	0.319	0.202	0.344	0.353	0.381
Montespertoli	0.762	0.554	0.553	0.735	0.653	0.638	0.678	0.682
Montevarchi	0.696	0.515	0.463	0.361	0.859	0.557	0.643	0.640
Monteverdi Marittimo	0.263	0.252	0.389	1.000	0.239	0.438	0.294	0.366
Monticiano	0.409	0.215	0.247	0.211	0.376	0.280	0.330	0.335
Montieri	0.258	0.044	0.291	0.206	0.396	0.236	0.259	0.272
Montignoso	0.608	0.285	0.508	0.160	0.429	0.394	0.440	0.470
Montopoli in Val d'Arno	0.887	0.483	0.524	0.469	0.469	0.550	0.639	0.647
Mulazzo	0.730	0.189	0.541	0.099	0.302	0.368	0.411	0.479
Murlo	0.474	1.000	0.337	0.502	0.225	0.508	0.425	0.439
Orbetello	0.455	0.327	0.342	0.389	0.350	0.367	0.395	0.394
Orciano Pisano	1.000	0.053	0.528	0.504	0.178	0.438	0.590	0.595
Ortignano Raggiolo	0.734	0.097	0.318	0.150	0.072	0.260	0.345	0.389
Palaia	0.589	0.333	0.397	0.263	0.377	0.383	0.449	0.448
Palazzuolo sul Senio	0.397	0.206	0.554	0.186	0.043	0.298	0.237	0.300
Peccioli	0.298	0.428	0.258	0.257	0.069	0.266	0.217	0.243
Pelago	0.649	0.379	0.337	0.384	0.333	0.402	0.468	0.470
Pergine Valdarno	0.794	0.321	0.467	0.309	0.389	0.442	0.568	0.549
Pescaglia	0.631	0.246	0.321	0.114	0.290	0.306	0.369	0.407
Pescia	0.695	0.435	0.429	0.285	0.783	0.503	0.591	0.601

Pian di Sco	0.813	0.582	0.545	0.344	0.360	0.522	0.605	0.588
Piancastagnaio	0.707	0.487	0.412	0.251	0.274	0.416	0.486	0.485
Piazza al Serchio	0.730	0.401	0.342	0.191	0.354	0.385	0.483	0.486
Pienza	0.464	0.207	0.262	0.184	0.316	0.275	0.347	0.342
Pietrasanta	0.428	0.274	0.424	0.259	0.503	0.374	0.408	0.412
Pieve a Nievole	0.985	0.382	0.391	0.544	0.376	0.507	0.590	0.643
Pieve Fosciana	0.672	0.476	0.486	0.416	0.420	0.487	0.571	0.536
Pieve Santo Stefano	0.571	0.837	0.393	0.193	0.068	0.416	0.285	0.400
Piombino	0.654	0.362	0.559	0.207	0.846	0.512	0.580	0.606
Pisa	0.553	0.555	0.608	0.442	0.723	0.575	0.578	0.588
Pistoia	0.784	0.858	0.487	0.490	0.763	0.656	0.677	0.700
Piteglio	0.351	0.177	0.234	0.137	0.400	0.249	0.291	0.304
Pitigliano	0.485	0.377	0.371	0.257	0.946	0.465	0.452	0.537
Podenzana	0.756	0.119	0.365	0.250	0.348	0.348	0.482	0.484
Poggibonsi	1.000	0.460	0.395	0.404	0.918	0.592	0.712	0.765
Poggio a Caiano	0.378	0.583	0.403	0.738	0.442	0.509	0.416	0.457
Pomarance	0.310	0.477	0.336	0.282	0.432	0.367	0.334	0.353
Ponsacco	0.893	0.647	0.439	0.562	0.951	0.663	0.737	0.771
Pontassieve	0.594	0.495	0.303	0.429	0.424	0.433	0.466	0.478
Ponte Buggianese	0.962	0.843	0.421	0.744	0.725	0.706	0.747	0.781
Pontedera	0.493	0.384	0.364	0.151	0.640	0.393	0.440	0.456
Pontremoli	0.457	0.494	0.339	0.273	0.558	0.413	0.431	0.441
Poppi	0.717	0.382	0.436	0.239	0.113	0.373	0.336	0.445
Porcari	0.653	0.509	0.452	0.431	0.276	0.460	0.485	0.493
Porto Azzurro	0.396	0.272	0.889	0.221	0.326	0.457	0.388	0.430
Portoferraio	0.498	0.432	0.471	0.254	0.411	0.413	0.439	0.438
Prato	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Pratovecchio	0.787	0.437	0.355	0.159	0.159	0.365	0.394	0.465
Quarrata	0.898	0.600	0.456	0.331	0.704	0.570	0.662	0.687
Radda in Chianti	0.387	0.276	0.437	0.298	0.426	0.367	0.371	0.385
Radicofani	0.441	0.076	0.402	0.118	0.292	0.265	0.340	0.331
Radicondoli	0.149	0.063	0.191	0.224	0.153	0.158	0.156	0.159
Rapolano Terme	0.495	0.371	0.425	0.321	0.379	0.396	0.437	0.425
Reggello	0.811	0.808	0.612	0.240	0.550	0.598	0.619	0.648
Rignano sull’Arno	0.854	0.624	0.365	0.706	0.325	0.553	0.567	0.613
Rio Marina	0.314	0.216	0.216	0.207	0.540	0.284	0.288	0.329
Rio nell’Elba	0.185	0.096	0.422	0.154	0.198	0.227	0.198	0.217
Riparbella	0.412	0.802	0.499	1.000	0.326	0.618	0.457	0.509
Roccalbegna	0.419	0.077	0.252	0.143	0.451	0.254	0.329	0.338
Roccastrada	0.685	0.518	0.295	0.400	0.995	0.541	0.587	0.645
Rosignano Marittimo	0.527	0.393	0.280	0.178	0.453	0.350	0.405	0.416
Rufina	0.780	0.277	0.465	0.509	0.262	0.448	0.506	0.534
Sambuca Pistoiese	0.426	0.446	0.390	0.116	1.000	0.457	0.369	0.518
San Casciano dei Bagni	0.404	0.234	0.181	0.217	0.267	0.247	0.309	0.299
San Casciano in Val di Pesa	0.807	0.459	0.371	0.286	0.765	0.506	0.596	0.634
San Gimignano	0.470	0.279	0.285	0.105	0.283	0.276	0.306	0.336
San Giovanni d’Asso	0.348	0.078	1.000	0.142	0.399	0.438	0.338	0.421
San Giovanni Valdarno	0.650	0.478	0.435	0.315	0.555	0.473	0.542	0.538
San Giuliano Terme	0.505	0.459	0.867	0.448	0.791	0.632	0.587	0.621
San Godenzo	0.437	1.000	0.483	0.087	0.216	0.458	0.331	0.396
San Marcello Pistoiese	0.624	0.426	0.289	0.210	0.600	0.405	0.468	0.497
San Miniato	0.624	0.519	0.596	0.428	0.614	0.554	0.593	0.585
San Piero a Sieve	0.708	0.507	0.331	0.309	0.266	0.409	0.475	0.478

San Quirico d'Orcia	0.530	0.337	0.410	0.336	0.437	0.403	0.462	0.450
San Romano in Garfagnana	0.453	0.249	0.673	0.104	0.101	0.341	0.313	0.349
San Vincenzo	0.297	0.359	0.420	0.264	0.225	0.324	0.293	0.302
Sansepolcro	0.745	0.578	0.418	0.264	0.919	0.556	0.637	0.660
Santa Croce sull'Arno	0.659	0.500	0.461	0.357	0.444	0.475	0.517	0.527
Santa Fiora	0.479	0.186	0.356	0.228	0.395	0.321	0.393	0.385
Santa Luce	0.473	0.839	0.260	0.432	0.215	0.438	0.394	0.402
Santa Maria a Monte	0.848	0.965	0.544	0.482	0.505	0.658	0.674	0.683
Sarteano	0.551	0.409	0.401	0.343	0.087	0.360	0.296	0.382
Sassetta	0.460	0.247	0.568	0.650	0.305	0.457	0.469	0.448
Scandicci	0.950	0.529	0.382	0.450	0.857	0.593	0.699	0.740
Scansano	0.410	0.296	0.343	0.254	0.518	0.356	0.387	0.396
Scarlino	0.348	0.246	0.433	0.244	0.227	0.309	0.323	0.313
Scarperia	0.516	0.522	0.359	0.279	0.355	0.400	0.430	0.424
Seggiano	0.380	0.073	0.365	0.175	0.344	0.266	0.335	0.320
Semproniano	0.506	0.073	0.201	0.226	0.283	0.240	0.354	0.335
Seravezza	0.649	0.463	0.398	0.427	0.581	0.487	0.556	0.550
Serravalle Pistoiese	0.837	0.565	0.687	0.427	0.386	0.581	0.620	0.636
Sestino	0.710	0.113	0.325	0.131	0.215	0.283	0.401	0.413
Sesto Fiorentino	0.827	0.360	0.766	0.514	0.564	0.607	0.650	0.681
Siena	0.406	0.329	0.256	0.335	0.523	0.355	0.394	0.394
Signa	0.742	0.677	0.566	0.454	0.659	0.609	0.658	0.653
Sillano	0.488	0.442	0.317	0.107	0.236	0.313	0.335	0.351
Sinalunga	0.703	0.758	0.487	0.390	0.834	0.615	0.657	0.664
Sorano	0.557	0.456	0.302	0.326	0.560	0.422	0.475	0.480
Sovicille	0.797	0.942	0.309	0.590	0.529	0.607	0.590	0.640
Stazzema	0.549	0.410	0.261	0.109	0.224	0.299	0.336	0.361
Stia	0.597	0.474	0.336	0.198	0.321	0.374	0.434	0.431
Subbiano	0.772	0.863	0.470	0.196	0.422	0.534	0.542	0.578
Suvereto	0.760	0.375	0.473	0.539	0.274	0.476	0.546	0.542
Talla	0.616	0.200	0.302	0.143	0.205	0.281	0.384	0.378
Tavarnelle Val di Pesa	0.585	0.715	0.563	0.160	0.371	0.484	0.454	0.491
Terranuova Bracciolini	0.448	0.443	0.467	0.404	0.398	0.435	0.436	0.434
Terricciola	0.603	0.403	0.493	0.217	0.311	0.405	0.457	0.454
Torrita di Siena	0.653	0.686	0.411	0.262	0.990	0.574	0.574	0.647
Trequanda	0.417	0.216	0.298	0.182	0.261	0.270	0.341	0.316
Tresana	0.551	0.426	0.445	0.304	0.437	0.428	0.487	0.467
Uzzano	0.802	0.602	0.515	0.648	0.334	0.572	0.625	0.611
Vagli Sotto	0.342	0.390	0.380	0.129	0.478	0.343	0.315	0.358
Vaglia	0.518	0.318	0.361	0.258	0.143	0.318	0.358	0.357
Vaiano	0.702	0.538	0.309	0.456	0.447	0.469	0.524	0.535
Vecchiano	0.742	0.370	0.561	0.460	0.592	0.534	0.613	0.613
Vernio	0.550	0.538	0.320	0.380	0.264	0.402	0.428	0.424
Viareggio	0.434	0.320	0.511	0.436	0.493	0.442	0.446	0.452
Vicchio	0.711	0.553	0.428	0.402	0.329	0.474	0.523	0.525
Vicopisano	0.979	0.450	0.531	0.295	0.460	0.523	0.630	0.659
Villa Basilica	0.213	0.114	0.265	0.186	0.194	0.198	0.211	0.206
Villa Collemandina	0.795	0.331	0.342	0.136	0.378	0.374	0.477	0.506
Villafranca in Lunigiana	0.599	0.214	0.350	0.229	0.328	0.332	0.417	0.419
Vinci	0.754	0.601	0.612	0.293	0.505	0.549	0.592	0.604
Volterra	0.565	0.567	0.286	0.282	0.509	0.424	0.467	0.472
Zeri	0.521	0.074	0.206	0.076	0.508	0.253	0.317	0.376

References

Aaronson, D. and Mazumder, B. (2011). The impact of rosenwald schools on black achievement. *Journal of Political Economy*, 119(5):821–888. 112

Adler, N. and Volta, N. (2016). Accounting for externalities and disposability: A directional economic environmental distance function. *European Journal of Operational Research*, 250(1):314–327. 10, 55, 57, 58

Afonso, A. and Fernandes, S. (2006). Measuring local government spending efficiency: evidence for the lisbon region. *Regional Studies*, 40(1):39–53. 84, 94

Afonso, A. and Fernandes, S. (2008). Assessing and explaining the relative efficiency of local government. *The Journal of Socio-Economics*, 37(5):1946–1979. 10, 28, 82, 84, 86, 93

Afonso, A., Schuknecht, L., and Tanzi, V. (2005). Public sector efficiency: an international comparison. *Public choice*, 123(3):321–347. 31, 86

Afonso, A. and Venâncio, A. (2016). The relevance of commuting zones for regional spending efficiency. *Applied Economics*, 48(10):865–877. 82, 84, 94

Agasisti, T., Dal Bianco, A., and Griffini, M. (2016). The public sector efficiency in italy: The case of lombardy municipalities in the provision of the essential public services. *Economia Pubblica*. 84

Agasisti, T. and Wolszczak-Derlacz, J. (2015). Exploring efficiency differentials between italian and polish universities, 2001–11. *Science and Public Policy*, 43(1):128–142. 92

Aigner, D., Lovell, C. K., and Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of econometrics*, 6(1):21–37. 8, 85

Akhmouch, A. and Correia, F. N. (2016). The 12 oecd principles on water governance when science meets policy. *Utilities Policy*, 43(Part A):14 – 20. 146

An, Q., Chen, H., Xiong, B., Wu, J., and Liang, L. (2016). Target intermediate products setting in a two-stage system with fairness concern. *Omega*. 62

Angrist, J. D., Battistin, E., and Vuri, D. (2014). In a small moment: Class size and moral hazard in the mezzogiorno. Working Paper 20173, National Bureau of Economic Research. 122

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press. 12

Antonini, E., Boeri, A., Gaspari, J., Gianfrate, V., and Longo, D. (2015). The quality of school buildings: a national emergency, a research field. *TECHNE: Journal of Technology for Architecture & Environment*, 9. 139

Asatryan, Z. and De Witte, K. (2015). Direct democracy and local government efficiency. *European Journal of Political Economy*, 39:58–66. 2, 10, 28, 141

Athanassopoulos, A. D. and Triantis, K. P. (1998). Assessing aggregate cost efficiency and the related policy implications for greek local municipalities. *INFOR: Information Systems and Operational Research*, 36(3):66–83. 84, 105

Bădin, L., Daraio, C., and Simar, L. (2012). How to measure the impact of environmental factors in a nonparametric production model. *European Journal of Operational Research*, 223(3):818–833. 10, 43

Bakó-Biró, Z., Clements-Croome, D., Kochhar, N., Awbi, H., and Williams, M. (2012). Ventilation rates in schools and pupils performance. *Building and Environment*, 48:215–223. 7, 113

Balaguer-Coll, M. T., Prior, D., and Tortosa-Ausina, E. (2007). On the determinants of local government performance: A two-stage nonparametric approach. *European Economic Review*, 51(2):425–451. 84, 86, 93, 148

Balaguer-Coll, M. T., Prior, D., and Tortosa-Ausina, E. (2013). Output complexity, environmental conditions, and the efficiency of municipalities. *Journal of productivity analysis*, 39(3):303–324. 84, 106

Banker, R. D. and Chang, H. (2006). The super-efficiency procedure for outlier identification, not for ranking efficient units. *European Journal of Operational Research*, 175(2):1311–1320. 94

Banker, R. D., Charnes, A., and Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science*, 30(9):1078–1092. 88

Barros, C. P., Managi, S., and Matousek, R. (2012). The technical efficiency of the Japanese banks: non-radial directional performance measurement with undesirable output. *Omega*, 40(1):1–8. 58

Battistin, E. and Meroni, E. C. (2016). Should we increase instruction time in low achieving schools? evidence from southern italy. *Economics of Education Review*, 55:39–56. 122

Benito, B., Bastida, F., and García, J. A. (2010). The determinants of efficiency in municipal governments. *Applied Economics*, 42(4):515–528. 84, 106

Berg, S. and Marques, R. (2011). Quantitative studies of water and sanitation utilities: a benchmarking literature survey. *Water Policy*, 13(5):591–606. 56

Bernini, C., Guizzardi, A., and Angelini, G. (2013). Dea-like model and common weights approach for the construction of a subjective community well-being indicator. *Social indicators research*, 114(2):405–424. 87, 89, 90, 150

Boetti, L., Piacenza, M., and Turati, G. (2012). Decentralization and local governments' performance: how does fiscal autonomy affect spending efficiency? *FinanzArchiv: Public Finance Analysis*, 68(3):269–302. 84, 93, 94, 105, 107

Bogetoft, P. and Otto, L. (2010). *Benchmarking with Dea, Sfa, and R*, volume 157. Springer Science & Business Media. 13, 106

Bollino, C. A., Di Vaio, G., and Polinori, P. (2012). Spillover ed eterogeneità spaziali nei livelli d'efficienza delle amministrazioni locali: Un'applicazione ai comuni dell'emilia-romagna. 84

Bönisch, P., Haug, P., Illy, A., and Schreier, L. (2011). Municipality size and efficiency of local public services: Does size matter? Technical report, IWH Discussion papers. 84, 85

Borge, L.-E., Falch, T., and Tovmo, P. (2008). Public sector efficiency: the roles of political and budgetary institutions, fiscal capacity, and democratic participation. *Public Choice*, 136(3):475–495. 84, 94

Bosch, N., Espasa, M., and Mora, T. (2012). Citizen control and the efficiency of local public services. *Environment and Planning C: Government and Policy*, 30(2):248–266. 10, 28, 39

Bowlin, W. F. (1998). Measuring performance: An introduction to data envelopment analysis (dea). *The Journal of Cost Analysis*, 15(2):3–27. 89

Branham, D. (2004). The wise man builds his house upon the rock: The effects of inadequate school building infrastructure on student attendance. *Social Science Quarterly*, 85(5):1112–1128. 113

Calabria, F. A., Camanho, A. S., and Zanella, A. (2016). The use of composite indicators to evaluate the performance of Brazilian hydropower plants. *International Transactions in Operational Research*. 38

Card, D. and Krueger, A. B. (1996). School resources and student outcomes: An overview of the literature and new evidence from north and south carolina. *The Journal of Economic Perspectives*, 10(4):31–50. 111, 112

Cazals, C., Fève, F., Florens, J.-P., and Simar, L. (2016). Nonparametric instrumental variables estimation for efficiency frontier. *Journal of Econometrics*, 190(2):349–359. 151

Cazals, C., Florens, J.-P., and Simar, L. (2002). Nonparametric frontier estimation: a robust approach. *Journal of econometrics*, 106(1):1–25. 41

Cellini, S. R., Ferreira, F., and Rothstein, J. (2010). The value of school facility investments: Evidence from a dynamic regression discontinuity design. *The Quarterly Journal of Economics*, 125(1):215–261. 112, 148, 149

Chambers, R. G., Chung, Y., and Färe, R. (1996). Benefit and distance functions. *Journal of economic theory*, 70(2):407–419. 58

Chambers, R. G., Chung, Y., and Färe, R. (1998). Profit, directional distance functions, and Nerlovian efficiency. *Journal of Optimization Theory and Applications*, 98(2):351–364. 58

Charnes, A., Cooper, W. W., and Rhodes, E. (1978). Measuring the efficiency of decision making units. *European journal of operational research*, 2(6):429–444. 8, 35, 85, 88

Chen, C.-M. and Ang, S. (2016). Measuring Environmental Efficiency: An Application to US Electric Utilities. In Zhu, J., editor, *Data Envelopment Analysis*, pages 345–366. Springer, New York. 60

Cheng, G. and Zervopoulos, P. D. (2014). Estimating the technical efficiency of health care systems: A cross-country comparison using the directional distance function. *European Journal of Operational Research*, 238(3):899–910. 58

Cherchye, L., Moesen, W., Rogge, N., and Van Puyenbroeck, T. (2007). An introduction to benefit of the doubt composite indicators. *Social indicators research*, 82(1):111–145. 9, 11, 36, 38, 39, 86, 89, 149

Chung, Y. H., Färe, R., and Grosskopf, S. (1997). Productivity and undesirable outputs: a directional distance function approach. *Journal of Environmental Management*, 51(3):229–240. 37, 58

Cipollone, P. and Rosolia, A. (2007). Social interactions in high school: Lessons from an earthquake. *The American economic review*, 97(3):948–965. 126

Coelli, T. (1996). A guide to deap version 2.1: a data envelopment analysis (computer) program. *Centre for Efficiency and Productivity Analysis, University of New England, Australia*. 95

Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., and Battese, G. E. (2005). *An introduction to efficiency and productivity analysis*. Springer Science & Business Media. 40

Coleman, J. S. (1966). Equality of educational opportunity. *Government Printing Office, Washington*, (4519). 7, 20

Cook, W. D., Tone, K., and Zhu, J. (2014). Data envelopment analysis: Prior to choosing a model. *Omega*, 44:1–4. 89

Cooper, W. W., Li, S., Seiford, L. M., and Zhu, J. (2011). Sensitivity analysis in dea. In *Handbook on Data Envelopment Analysis*, pages 71–91. Springer. 88, 89

Cordero, J. M., Pedraja-Chaparro, F., Pisaflores, E. C., and Polo, C. (2016). Efficiency assessment of Portuguese municipalities using a conditional nonparametric approach. *Journal of Productivity Analysis*, pages 1–24. 10, 27, 28, 42, 47, 149, 151

Cordero, J. M., Salinas-Jiménez, J., and Salinas-Jiménez, M. M. (2017). Exploring factors affecting the level of happiness across countries: A conditional robust nonparametric frontier analysis. *European Journal of Operational Research*, 256(2):663–672. 41

Cordero, J. M., Santín, D., and Sicilia, G. (2015). Testing the accuracy of dea estimates under endogeneity through a monte carlo simulation. *European Journal of Operational Research*, 244(2):511–518. 151

Cuadrado-Ballesteros, B., García-Sánchez, I.-M., and Prado-Lorenzo, J.-M. (2013). Effect of modes of public services delivery on the efficiency of local governments: A two-stage approach. *Utilities Policy*, 26:23–35. 84, 106

Da Cruz, N. F. and Marques, R. C. (2014). Revisiting the determinants of local government performance. *Omega*, 44:91–103. 1, 4, 82, 84, 86, 92, 93, 94, 105, 106, 141

Da Cruz, N. F., Marques, R. C., Romano, G., and Guerrini, A. (2012). Measuring the efficiency of water utilities: a cross-national comparison between Portugal and Italy. *Water Policy*, 14(5):841–853. 66

Dakpo, K. H., Jeanneaux, P., and Latruffe, L. (2016). Modelling pollution-generating technologies in performance benchmarking: recent developments, limits and future prospects in the nonparametric framework. *European Journal of Operational Research*, 250(2):347–359. 37, 57, 58, 67

Daraio, C. and Simar, L. (2005). Introducing environmental variables in nonparametric frontier models: a probabilistic approach. *Journal of Productivity Analysis*, 24(1):93–121. 41

Daraio, C. and Simar, L. (2007). *Advanced robust and nonparametric methods in efficiency analysis: Methodology and applications*. Springer Science & Business Media. 10, 41, 43

Daraio, C., Simar, L., and Wilson, P. W. (2017). Central limit theorems for conditional efficiency measures and tests of the separability condition in nonparametric, two-stage models of production. *The Econometrics Journal*. 13, 92

De Borger, B. and Kerstens, K. (1996). Cost efficiency of belgian local governments: A comparative analysis of fdh, dea, and econometric approaches. *Regional science and urban economics*, 26(2):145–170. 84, 85, 86, 93, 94, 105

De Borger, B. and Kerstens, K. (2000). What is known about municipal efficiency? in public provision and performance: Contributions from efficiency and productivity measurement. In *Blank, J.L.T. (ed.), Public provision and performance - contributions from efficiency and productivity measurement*, pages 299–330. Elsevier, Amsterdam et al. 84

De Borger, B., Kerstens, K., Moesen, W., and Vanneste, J. (1994). Explaining differences in productive efficiency: An application to belgian municipalities. *Public Choice*, 80(3):339–358. 84, 86, 94

De Sousa, M. D. C. S. and Stošić, B. (2005). Technical efficiency of the brazilian municipalities: correcting nonparametric frontier measurements for outliers. *Journal of Productivity analysis*, 24(2):157–181. 84, 93, 106

De Witte, K. and Geys, B. (2009). The political economy of efficient public good provision: evidence from Flemish libraries using a generalised conditional efficiency framework. 34, 153

De Witte, K. and Geys, B. (2011). Evaluating efficient public good provision: Theory and evidence from a generalised conditional efficiency model for public libraries. *Journal of urban economics*, 69(3):319–327. 32

De Witte, K. and Geys, B. (2013). Citizen coproduction and efficient public good provision: Theory and evidence from local public libraries. *European Journal of Operational Research*, 224(3):592–602. 32, 150

De Witte, K. and Kortelainen, M. (2013). What explains the performance of students in a heterogeneous environment? Conditional efficiency estimation with continuous and discrete environmental variables. *Applied Economics*, 45(17):2401–2412. 41

De Witte, K. and López-Torres, L. (2017). Efficiency in education: a review of literature and a way forward. *Journal of the Operational Research Society*, 68(4):339–363. 12, 151

De Witte, K. and Marques, R. C. (2010). Influential observations in frontier models, a robust non-oriented approach to the water sector. *Annals of Operations Research*, 181(1):377–392. 55, 57, 66

De Witte, K. and Marques, R. C. (2012). Gaming in a benchmarking environment. a non-parametric analysis of benchmarking in the water sector. *Water Policy*, 14(1):45–66. 66

De Witte, K. and Moesen, W. (2010). Sizing the government. *Public Choice*, 145(1):39–55. 5, 27, 146

Deller, S. C., Chicoine, D. L., and Walzer, N. (1988). Economies of size and scope in rural low-volume roads. *The Review of Economics and Statistics*, pages 459–465. 85

Deprins, D., Simar, L., and Tulkens, H. (1984). Measuring labour-efficiency in post offices. In *The Performance of Public Enterprises: Concepts and Measurement, North-Holland*, pages 243–267. Amsterdam: North Holland. 85

Despotis, D. (2005). A reassessment of the human development index via data envelopment analysis. *Journal of the Operational Research Society*, 56(8):969–980. 87

Despotis, D. K. et al. (2002). Improving the discriminating power of dea: Focus on globally efficient units. *Journal of the Operational Research Society*, 53(3):314–323. 12, 87, 90

D'Inverno, G., Carosi, L., and Ravagli, L. (2017). Global public spending efficiency in Tuscan municipalities. *Socio-Economic Planning Sciences*. 28, 39

Dollery, B. and Wallis, J. L. (2001). *The political economy of local government*. Edward Elgar Publishing. 85, 86

Doumpos, M. and Cohen, S. (2014). Applying data envelopment analysis on accounting data to assess and optimize the efficiency of greek local governments. *Omega*, 46:74–85. 82, 84, 85, 93, 105, 106, 148

Dunleavy, P., Margetts, H., Bastow, S., and Tinkler, J. (2006). New public management is deadlong live digital-era governance. *Journal of public administration research and theory*, 16(3):467–494. 82

Dyson, R. G., Allen, R., Camanho, A. S., Podinovski, V. V., Sarrico, C. S., and Shale, E. A. (2001). Pitfalls and protocols in dea. *European Journal of operational research*, 132(2):245–259. 89

Earthman, G. I. (2002). School facility conditions and student academic achievement. *UCLA's Institute for Democracy, Education, & Access*. 7, 113

Eeckaut, P., Tulkens, H., and Jamar, M.-A. (1993). Cost efficiency in belgian municipalities. In *Fried H, Lovell C, Schmidt S, editors. The Measurement of Productive Efficiency–Techniques and Applications*, pages 300–334. Oxford University Press New York, NY. 85

European Commission (2007). Terms and Definitions of the Urban Waste Water Treatment Directive 91/271/EEC, Brussels. 70

European Commission (2014). Financing Schools in Europe: Mechanisms, Methods and Criteria in Public Funding. *Brussels, Education, Audiovisual and Culture Executive Agency (EACEA)/Eurydice Report*. 7, 116, 117

Färe, R. and Grosskopf, S. (2003). *New directions: efficiency and productivity*. Kluwer Academic Publishers, Boston. 59

Färe, R. and Grosskopf, S. (2004). Modeling undesirable factors in efficiency evaluation: comment. *European Journal of Operational Research*, 157(1):242–245. 58

Färe, R. and Grosskopf, S. (2010). Directional distance functions and slacks-based measures of efficiency. *European Journal of Operational Research*, 200(1):320–322. 58

Färe, R., Grosskopf, S., Lovell, C. K., and Pasurka, C. (1989). Multilateral productivity comparisons when some outputs are undesirable: a nonparametric approach. *The review of Economics and Statistics*, pages 90–98. 57, 60

Färe, R., Grosskopf, S., and Pasurka, C. (2016). Technical change and pollution abatement costs. *European Journal of Operational Research*, 248(2):715–724. 57

Färe, R., Grosskopf, S., and Pasurka, C. A. (2014). Potential gains from trading bad outputs: The case of us electric power plants. *Resource and Energy Economics*, 36(1):99–112. 57, 60, 67

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society. Series A (General)*, 120(3):253–290. 8, 84, 85

Førsund, F. R. (2008). Good modelling of bad outputs: pollution and multiple-output production. Technical report, Memorandum, Department of Economics, University of Oslo. 58

Fuentes, R., Torregrosa, T., and Ballenilla, E. (2015). Conditional order-m efficiency of wastewater treatment plants: The role of environmental factors. *Water*, 7(10):5503–5524. 56, 65, 66

Fukuyama, H. and Weber, W. L. (2009). A directional slacks-based measure of technical inefficiency. *Socio-Economic Planning Sciences*, 43(4):274–287. 58

Fusco, E. (2015). Enhancing non-compensatory composite indicators: a directional proposal. *European Journal of Operational Research*, 242(2):620–630. 40

Fusco, E., Vidoli, F., and Sahoo, B. K. (2017). Spatial heterogeneity in composite indicator: A methodological proposal. *Omega*. 10, 27

Galli, P., Castenetto, S., and Peronace, E. (2012). The mcs macroseismic survey of the emilia 2012 earthquakes. *Annals of Geophysics*, 55(4). 123

Geys, B., Heinemann, F., and Kalb, A. (2008). Local governments in the wake of demographic change: Evidence from german municipalities. *FinanzArchiv: Public Finance Analysis*, 64(4):434–457. 148

Geys, B., Heinemann, F., and Kalb, A. (2010). Voter involvement, fiscal autonomy and public sector efficiency: evidence from german municipalities. *European journal of political economy*, 26(2):265–278. 84, 93, 94, 106

Geys, B. and Moesen, W. (2009). Measuring local government technical (in) efficiency: an application and comparison of fdh, dea, and econometric approaches. *Public Performance & Management Review*, 32(4):499–513. 84, 94

Global Taskforce of Local and Regional Governments (2016). Roadmap for localizing the SDGs: Implementation and monitoring at subnational level. 1, 2

Golany, B., Roll, Y., and Rybak, D. (1994). Measuring efficiency of power plants in Israel by data envelopment analysis. *IEEE Transactions on Engineering Management*, 41(3):291–301. 66

Goodland, R. (1995). The concept of environmental sustainability. *Annual review of ecology and systematics*, 26(1):1–24. 54

Greenwald, R., Hedges, L. V., and Laine, R. D. (1996). The effect of school resources on student achievement. *Review of educational research*, 66(3):361–396. 111

Grünthal, G. (1998). European macroseismic scale 1998. Technical report, European Seismological Commission (ESC). 120

Grupp, H. and Schubert, T. (2010). Review and new evidence on composite innovation indicators for evaluating national performance. *Research Policy*, 39(1):67–78. 40

Guerrini, A., Romano, G., Mancuso, F., and Carosi, L. (2016). Identifying the performance drivers of wastewater treatment plants through conditional order-m efficiency analysis. *Utilities Policy*, 42:20–31. 72, 147

Hampf, B. and Rødseth, K. L. (2015). Carbon dioxide emission standards for us power plants: An efficiency analysis perspective. *Energy Economics*, 50:140–153. 57

Hanushek, E. A. (1996). Measuring investment in education. *The Journal of Economic Perspectives*, 10(4):9–30. 112

Haverinen-Shaughnessy, U., Shaughnessy, R. J., Cole, E. C., Toyinbo, O., and Moschandreas, D. J. (2015). An assessment of indoor environmental quality in schools and its association with health and performance. *Building and Environment*, 93:35–40. 7, 113

Hayes, K. and Chang, S. (1990). The relative efficiency of city manager and mayor-council forms of government. *Southern Economic Journal*, pages 167–177. 85

Helland, L. and Sørensen, R. J. (2015). Partisan bias, electoral volatility, and government efficiency. *Electoral Studies*, 39:117–128. 10, 28, 39

Hermenier, H. (2009). The Quest for "Perfect Territorial Organization": Comparison across Europe. In Council of European Municipalities and Regions (CEMR), editor, *Local and regional authorities. Balancing democracy, identity and efficiency - Changes in local and regional structures in Europe*, pages 4–7. CEMR, Brussel. 6

Hernández-Sancho, F., Molinos-Senante, M., and Sala-Garrido, R. (2011a). Energy efficiency in Spanish wastewater treatment plants: a non-radial DEA approach. *Science of the Total Environment*, 409(14):2693–2699. 56, 66, 70

Hernández-Sancho, F., Molinos-Senante, M., and Sala-Garrido, R. (2011b). Techno-economical efficiency and productivity change of wastewater treatment plants: the role of internal and external factors. *Journal of Environmental Monitoring*, 13(12):3448–3459. 56, 64, 65, 66, 70, 71, 73, 147

Hernández-Sancho, F., Molinos-Senante, M., Sala-Garrido, R., and Del Saz-Salazar, S. (2012). Tariffs and efficient performance by water suppliers: an empirical approach. *Water Policy*, 14(5):854–864. 55, 57

Hernández-Sancho, F. and Sala-Garrido, R. (2009). Technical efficiency and cost analysis in wastewater treatment processes: A DEA approach. *Desalination*, 249(1):230–234. 56, 66

Hood, C. (1991). A public management for all seasons? *Public administration*, 69(1):3–19. 6, 82

Hsiao, C., Yang, C., and Bjornlund, H. (2007). Performance measurement in wastewater control-pig farms in Taiwan. *WIT Transactions on Ecology and the Environment*, 103:467–474. 66

Hughes, O. E. (2003). Public management and administration. 82

Hyndman, N. and Lapsley, I. (2016). New public management: The story continues. *Financial Accountability & Management*, 32(4):385–408. 6

Iommi, S. (2011). Numerosità e dimensione degli enti locali ed offerta dei servizi pubblici. *IRPET, Il governo locale in Toscana. Identikit*. 83

Jackson, C. K., Johnson, R. C., and Persico, C. (2016). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *The Quarterly Journal of Economics*, 157:218. 111, 112

Jones, J. T. and Zimmer, R. W. (2001). Examining the impact of capital on academic achievement. *Economics of Education Review*, 20(6):577–588. 112

Kalb, A., Geys, B., and Heinemann, F. (2012). Value for money? german local government efficiency in a comparative perspective. *Applied economics*, 44(2):201–218. 84, 93, 94, 106

Karagiannis, G. (2017). On aggregate composite indicators. *Journal of the Operational Research Society*, pages 1–6. 4, 10, 28

Keating, M. (1998). *Public management reform and economic and social development*. OECD. 82

Kornai, J., Maskin, E., and Roland, G. (2003). Understanding the soft budget constraint. *Journal of economic literature*, 41(4):1095–1136. 105, 145

Kruskal, W. H. and Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, 47(260):583–621. 14, 65

Kuosmanen, T. (2005). Weak disposability in nonparametric production analysis with undesirable outputs. *American Journal of Agricultural Economics*, 87(4):1077–1082. 55, 60

Lawrence, B. K. (2003). Save a penny, lose a school: The real cost of deferred maintenance. policy brief series on rural education. 113

Liu, J. S., Lu, L. Y., and Lu, W.-M. (2016). Research fronts in data envelopment analysis. *Omega*, 58:33–45. 58

Lo Storto, C. (2013). Evaluating technical efficiency of italian major municipalities: a data envelopment analysis model. *Procedia-Social and Behavioral Sciences*, 81:346–350. 84

Lo Storto, C. (2016). The trade-off between cost efficiency and public service quality: A non-parametric frontier analysis of italian major municipalities. *Cities*, 51:52–63. 82, 84, 92, 93, 105, 106

Loikkanen, H. A. and Susiluoto, I. (2005). Cost efficiency of finnish municipalities in basic service provision 1994-2002. *Urban Public Economics Review*, (4). 84

Lorenzo-Toja, Y., Vázquez-Rowe, I., Chenel, S., Marín-Navarro, D., Moreira, M. T., and Feijoo, G. (2015). Eco-efficiency analysis of Spanish WWTPs using the LCA+ DEA method. *Water Research*, 68:651–666. 67

Martorell, P., Stange, K., and McFarlin, I. (2016). Investing in schools: capital spending, facility conditions, and student achievement. *Journal of Public Economics*, 140:13–29. 112, 114, 148, 149

Mastromarco, C. and Simar, L. (2015). Effect of FDI and time on catching up: new insights from a conditional nonparametric frontier analysis. *Journal of Applied Econometrics*, 30(5):826–847. 10, 27, 42

Melyn, W. and Moesen, W. (1991). Towards a synthetic indicator of macroeconomic performance: unequal weighting when limited information is available. 9, 11, 35, 149

Mendell, M. J. and Heath, G. A. (2004). Do indoor environments in schools influence student performance? a review of the literature. *Lawrence Berkeley National Laboratory*. 7, 113

Meroni, F., Squarcina, T., Pessina, V., Locati, M., Modica, M., and Zoboli, R. (2017). A damage scenario for the 2012 northern italy earthquakes and estimation of the economic losses to residential buildings. *International Journal of Disaster Risk Science*. 123

Metcalf & Eddy, I., Tchobanoglous, G., Burton, F., and Stensel, H. D. (2003). *Wastewater Engineering: Treatment and Reuse*. McGraw-Hill Education, New York, 4 edition. 147, 179, 180, 181

Moldan, B., Janoušková, S., and Hák, T. (2012). How to understand and measure environmental sustainability: Indicators and targets. *Ecological Indicators*, 17:4–13. 54

Molinos-Senante, M., Hernández-Sancho, F., Mocholí-Arce, M., and Sala-Garrido, R. (2014a). Economic and environmental performance of wastewater treatment plants: potential reductions in greenhouse gases emissions. *Resource and Energy Economics*, 38:125–140. 55, 56, 57, 58, 66, 70, 72, 73, 146, 147

Molinos-Senante, M., Hernandez-Sancho, F., and Sala-Garrido, R. (2014b). Benchmarking in wastewater treatment plants: a tool to save operational costs. *Clean Technologies and Environmental Policy*, 16(1):149–161. 56, 65, 66, 70, 72, 73, 147

Molinos-Senante, M., Hernández-Sancho, F., and Sala-Garrido, R. (2015a). Comparing the dynamic performance of wastewater treatment systems: A metafrontier Malmquist productivity index approach. *Journal of environmental management*, 161:309–316. 56, 66, 70

Molinos-Senante, M., Maziotis, A., and Sala-Garrido, R. (2016). Estimating the cost of improving service quality in water supply: a shadow price approach for England and Wales. *Science of The Total Environment*, 539:470–477. 57, 58

Molinos-Senante, M., Sala-Garrido, R., and Lafuente, M. (2015b). The role of environmental variables on the efficiency of water and sewerage companies: a case study of Chile. *Environmental Science and Pollution Research*, 22(13):10242–10253. 55, 57, 58

Morais, P. and Camanho, A. S. (2011). Evaluation of performance of european cities with the aim to promote quality of life improvements. *Omega*, 39(4):398–409. 31

Murty, S. (2010). On the theory of a firm: the case of by-production of emissions. Technical report, University of Warwick. Dept. of Economics. 57, 67

Murty, S., Russell, R. R., and Levkoff, S. B. (2012). On modeling pollution-generating technologies. *Journal of Environmental Economics and Management*, 64(1):117–135. 67

Nakazawa, K. (2013). Cost inefficiency of municipalities after amalgamation. *Procedia Economics and Finance*, 5:581–588. 84, 93, 94

Narbón-Perpiñá, I. and De Witte, K. (2017a). Local governments' efficiency: A systematic literature review-Part I. *International Transactions in Operational Research*, pages 1–38. 27, 30, 32

Narbón-Perpiñá, I. and De Witte, K. (2017b). Local governments' efficiency: A systematic literature review-Part II. *International Transactions in Operational Research*, pages 1–30. 27, 33, 46

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., and Giovannini, E. (2005). Handbook on constructing composite indicators. 4, 9

Neilson, C. A. and Zimmerman, S. D. (2014). The effect of school construction on test scores, school enrollment, and home prices. *Journal of Public Economics*, 120:18–31. 111, 112

Nijkamp, P. and Suzuki, S. (2009). A generalized goals-achievement model in data envelopment analysis: an application to efficiency improvement in local government finance in japan. *Spatial economic analysis*, 4(3):249–274. 84

OECD (2001). OECD environmental strategy for the first decade of the 21st century. *Paris*. 54

OECD (2008). *Handbook on constructing composite indicators: methodology and user guide*. 4, 9, 28, 32, 143

OECD (2014). Recommendation of the Council on Effective Public Investment Across Levels of Government. *OECD Publishing, Paris*, 12 March 2014. 16

OECD (2015). OECD principles on water governance welcomed by ministers at the OECD ministerial council meeting on 4 june 2015. *Directorate for Public Governance and Territorial Development, Organization for Economic Cooperation and Development*. 147

OECD (2016a). OECD regions at a Glance 2016. *OECD Publishing, Paris*. 2

OECD (2016b). PISA 2015 results (Volume I). *OECD Publishing*. 148

OECD (2017a). Diffuse pollution, degraded waters: Emerging policy solutions. *OECD Studies on Water*. 146

OECD (2017b). Education at a Glance 2017. *OECD Publishing, Paris*. 17, 149

OECD (2017c). Government at a Glance 2017. *OECD Publishing, Paris*. 2, 13

OECD (2017d). Subnational governments in OECD countries: Key data (brochure). *OECD Publishing, Paris*. 15, 142

Osborne, S. P. (2006). The new public governance? *Public Management Review*, 8(3):377–387. 82

Pakkar, M. S. (2015). An integrated approach based on DEA and AHP. *Computational Management Science*, 12(1):153–169. 11, 59

Parks, R. B., Baker, P. C., Kiser, L., Oakerson, R., Ostrom, E., Ostrom, V., Percy, S. L., Vandivort, M. B., Whitaker, G. P., and Wilson, R. (1981). Consumers as coproducers of public services: Some economic and institutional considerations. *Policy Studies Journal*, 9(7):1001–1011. 150

Picazo-Tadeo, A. J., Reig-Martinez, E., and Hernandez-Sancho, F. (2005). Directional distance functions and environmental regulation. *Resource and Energy Economics*, 27(2):131–142. 58

Picazo-Tadeo, A. J., Sáez-Fernández, F. J., and González-Gómez, F. (2008). Does service quality matter in measuring the performance of water utilities? *Utilities Policy*, 16(1):30–38. 55, 57, 58

Picazo-Tadeo, A. J., Sáez-Fernández, F. J., and González-Gómez, F. (2011). Assessing performance in the management of the urban water cycle. *Water Policy*, 13(6):782–796. 58, 66

Pomerol, J.-C. and Barba-Romero, S. (2000). *Multicriterion decision in management: principles and practice*. Kluwer Academic Publishers, New York. 59, 62

Poveda, A. C. (2012). Estimating effectiveness of the control of violence and socioeconomic development in colombia: an application of dynamic data envelopment analysis and data panel approach. *Social indicators research*, 105(3):343–366. 105

Poveda, A. C. (2014). Corruption, economic development, and insecurity in colombia. *Handbook of research on strategic performance management and measurement using data envelopment analysis*, pages 373–387. 105

Prieto, A. M. and ZofIo, J. L. (2001). Evaluating effectiveness in public provision of infrastructure and equipment: the case of spanish municipalities. *Journal of productivity Analysis*, 15(1):41–58. 84

Reingewertz, Y. (2012). Do municipal amalgamations work? evidence from municipalities in israel. *Journal of Urban Economics*, 72(2):240–251. 105

Rogge, N. (2018). On aggregating benefit of the doubt composite indicators. *European Journal of Operational Research*, 264(1):364–369. 150

Rogge, N., De Jaeger, S., and Lavigne, C. (2017). Waste Performance of NUTS 2-regions in the EU: A Conditional Directional Distance Benefit-of-the-Doubt Model. *Ecological Economics*, 139:19–32. 10, 28, 34, 36, 37, 153

Rogoff, K. and Sibert, A. (1988). Elections and macroeconomic policy cycles. *The Review of Economic Studies*, 55(1):1–16. 107

Romano, G., Molinos-Senante, M., and Guerrini, A. (2017). Water utility efficiency assessment in italy by accounting for service quality: An empirical investigation. *Utilities Policy*, 45:97–108. 57, 150

Ruxton, G. D. and Beauchamp, G. (2008). Some suggestions about appropriate use of the Kruskal–Wallis test. *Animal behaviour*, 76(3):1083–1087. 14, 65

Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, 15(3):234–281. 59, 63

Saaty, T. L. (1990). How to make a decision: the analytic hierarchy process. *European Journal of Operational Research*, 48(1):9–26. 59, 63, 64

Sadioglu, U. (2016). *Theoretical Foundations and Discussions on the Reformation Process in Local Governments*. IGI Global. 15, 47

Sala-Garrido, R., Hernández-Sancho, F., and Molinos-Senante, M. (2012a). Assessing the efficiency of wastewater treatment plants in an uncertain context: a DEA with tolerances approach. *Environmental Science & Policy*, 18:34–44. 56, 66

Sala-Garrido, R., Molinos-Senante, M., and Hernández-Sancho, F. (2011). Comparing the efficiency of wastewater treatment technologies through a DEA metafrontier model. *Chemical Engineering Journal*, 173(3):766–772. 56, 66

Sala-Garrido, R., Molinos-Senante, M., and Hernández-Sancho, F. (2012b). How does seasonality affect water reuse possibilities? An efficiency and cost analysis. *Resources, Conservation and Recycling*, 58:125–131. 66

Santín, D. and Sicilia, G. (2017). Dealing with endogeneity in data envelopment analysis applications. *Expert Systems with Applications*, 68:173–184. 151

Sarrico, C. S. and Dyson, R. (2004). Restricting virtual weights in data envelopment analysis. *European Journal of Operational Research*, 159(1):17–34. 38

Simar, L., Vanhems, A., and Van Keilegom, I. (2016). Unobserved heterogeneity and endogeneity in nonparametric frontier estimation. *Journal of Econometrics*, 190(2):360–373. 151

Simar, L. and Wilson, P. W. (2000). A general methodology for bootstrapping in non-parametric frontier models. *Journal of applied statistics*, 27(6):779–802. 92

Simar, L. and Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, 136(1):31–64. 13, 65, 92

Slack, N. E. and Bird, R. M. (2013). *Merging municipalities: Is bigger better?* Institute on Municipal Finance & Governance, Munk School of Global Affairs, University of Toronto. 1, 141, 146, 148

Šťastná, L. and Gregor, M. (2015). Public sector efficiency in transition and beyond: evidence from czech local governments. *Applied Economics*, 47(7):680–699. 84, 105, 145

Sueyoshi, T. and Goto, M. (2012a). Data envelopment analysis for environmental assessment: comparison between public and private ownership in petroleum industry. *European journal of operational research*, 216(3):668–678. 58

Sueyoshi, T. and Goto, M. (2012b). Dea environmental assessment of coal fired power plants: Methodological comparison between radial and non-radial models. *Energy Economics*, 34(6):1854–1863. 58

Team, R. C. (2015). R: A language and environment for statistical computing [computer software]. r foundation for statistical computing, vienna, austria. 106

Tofallis, C. (2013). An automatic-democratic approach to weight setting for the new human development index. *Journal of Population Economics*, 26(4):1325–1345. 150

UN General Assembly (2002). Report of the world summit on sustainable development, johannesburg, south africa, 26 August–4 September 2002. *New York: United Nations*. 54

UN General Assembly (2015). Transforming our world: the 2030 agenda for sustainable development. *New York: United Nations*. 1, 54, 142

UN General Assembly (2017). New Urban Agenda. *New York: United Nations*. 2

United Nations High Level Political Forum (2017). PATHWAYS TO SUSTAINABLE DEVELOPMENT. First Belgian National Voluntary Review on the Implementation of the 2030 Agenda. *New York*. 15

Van Puyenbroeck, T. and Rogge, N. (2017). Geometric mean quantity index numbers with benefit-of-the-doubt weights. *European Journal of Operational Research*, 256(3):1004–1014. 40, 44, 150

Vannoli, P., Burrato, P., and Valensise, G. (2015). The seismotectonics of the po plain (northern italy): Tectonic diversity in a blind faulting domain. *Pure and Applied Geophysics*, 172(5):1105–1142. 119

Verbunt, P. and Rogge, N. (2018). Geometric composite indicators with compromise benefit-of-the-doubt weights. *European Journal of Operational Research*, 264(1):388–401. 150

Verschelde, M. and Rogge, N. (2012). An environment-adjusted evaluation of citizen satisfaction with local police effectiveness: Evidence from a conditional Data Envelopment Analysis approach. *European Journal of Operational Research*, 223(1):214–225. 36

Vidoli, F., Fusco, E., and Mazziotta, C. (2015). Non-compensability in composite indicators: A robust directional frontier method. *Social Indicators Research*, 122(3):635–652. 106

Vidoli, F. and Mazziotta, C. (2013). Robust weighted composite indicators by means of frontier methods with an application to european infrastructure endowment. *Italian journal of applied statistics*, 23(2):259–282. 106

Whitaker, G. P. (1980). Coproduction: Citizen participation in service delivery. *Public administration review*, pages 240–246. 150

Wilson, J. Q. and Kelling, G. L. (1982). Broken windows. *Atlantic monthly*, 249(3):29–38. 113

Worthington, A., Dollery, B., et al. (2000). *Efficiency measurement in the local public sector: Econometric and mathematical programming frontier techniques*. Queensland University of Technology, School of Economics and Finance. 84

Worthington, A. C. (2000). Cost efficiency in australian local government: A comparative analysis of mathematical programming and econometrical approaches. *Financial Accountability & Management*, 16(3):201–223. 85, 86, 105

Worthington, A. C. (2014). A review of frontier approaches to efficiency and productivity measurement in urban water utilities. *Urban Water Journal*, 11(1):55–73. 56

Yusfany, A. (2015). The Efficiency of Local Governments and its Influence Factors. *International Journal of Technology Enhancements and Emerging Engineering Research*, 4(10):219–241. 10, 28

Zanella, A., Camanho, A., and Dias, T. G. (2013). Benchmarking countries' environmental performance. *Journal of the Operational Research Society*, 64(3):426–438. 12

Zanella, A., Camanho, A. S., and Dias, T. G. (2015). Undesirable outputs and weighting schemes in composite indicators based on data envelopment analysis. *European Journal of Operational Research*, 245(2):517–530. 9, 26, 37, 38, 39

Zhang, N. and Choi, Y. (2014). A note on the evolution of directional distance function and its development in energy and environmental studies 1997–2013. *Renewable and Sustainable Energy Reviews*, 33:50–59. 58

Zhou, P., Ang, B., and Wang, H. (2012). Energy and CO_2 emission performance in electricity generation: a non-radial directional distance function approach. *European Journal of Operational Research*, 221(3):625–635. 10, 55, 58, 59, 61, 63, 64, 68, 69, 76, 182

Zhou, P., Ang, B. W., and Zhou, D. Q. (2010). Weighting and aggregation in composite indicator construction: a multiplicative optimization approach. *Social Indicators Research*, 96(1):169–181. 9