IMT School for Advanced Studies, Lucca

Lucca, Italy

Essays on Tax Collection and Local Government Efficiency

PhD Program in Economics

XXVIII Cycle

By

Duccio Gamannossi degl'Innocenti

2017

The dissertation of Duccio Gamannossi degl'Innocenti is approved.

Program Coordinator:
Prof. Massimo Riccaboni, IMT School for Advanced Studies, Lucca, Italy

Advisor:
Prof. Massimo Riccaboni, IMT School for Advanced Studies, Lucca, Italy

Co-Advisor:
Prof. Matthew D. Rablen, University of Sheffield, Sheffield, UK

Co-Advisor:
Prof. Andrea Vindigni, Università degli Studi di Genova, Genova, Italy

The dissertation of Duccio Gamannossi degl'Innocenti has been reviewed by:

Prof. Cinzia Daraio, Università di Roma La Sapienza, Rome, Italy

Prof. Roberto Dell'Anno, Università degli Studi di Salerno, Salerno, Italy

IMT School for Advanced Studies, Lucca

2017

A Fabiana

Contents

Acknowledgements

1. **Chapter** 2 was carried out in collaboration with Matthew D. Rablen, Reader at the University of Sheffield, and published on *Public Finance Review* on November 2016.
2. **Chapter** 3 was carried out in collaboration with Matthew D. Rablen, Reader at the University of Sheffield, and under submission as a contribution to the Routledge Companion on Tax Avoidance Research.
3. **Chapter** 4 was carried out in collaboration with Dr. Andrea Flori, and about to be submitted to the *Journal of Productivity Analysis*.

My sincere thanks to my Advisor, Prof. Riccaboni, for giving support when it was needed the most.
I am deeply in debt with my Co-Advisor, Prof. Matthew D. Rablen, who provided an endless and valuable source of inspiration and stimuli that contributed significantly to my own growth as a researcher.
I want to thank my Co-Advisor, Prof. Andrea Vindigni, I greatly appreciate the freedom you have given me to find my own path and the guidance and useful advices you offered when needed.
I am grateful to Prof. Alessandro Petretto for his invaluable support.
I also want to thank the two referees, Prof. Cinzia Daraio, whose works and insights guided me into the world efficiency analysis, and Prof. Roberto Dell'Anno, for his helpful and insightful comments.

Vita

October 3, 1985	Prato, Italy
2010	B.Sc. in Economics Final mark: 105/110 Università degli Studi di Firenze, Florence, Italy
2013	M.Sc. in Economics Final mark: 110/110 cum laude Università degli Studi di Firenze, Florence, Italy

Publications

1. D. Gamannossi degl'Innocenti, M. D. Rablen, "Income Tax Avoidance and Evasion. A Narrow Bracketing Approach, Public Finance Review, 2016

Presentations

1. D. Gamannossi degl'Innocenti and Andrea Flori "Italian Municipalities Efficiency and the Role of Exogenous Factors: A Conditional Frontier Model Approach." at *University of Rome La Sapienza* - Seminar for Course in Productivity and Efficiency Analysis, Rome, Italy, 2016
2. D. Gamannossi degl'Innocenti and Matthew D. Rablen "Optimal Income Tax Enforcement in the Presence of Tax Avoidance." at *Annual Workshop of the Tax Administration Research Centre - University of Exeter*, Exeter, UK, 2016
3. D. Gamannossi degl'Innocenti and Matthew D. Rablen "Optimal Income Tax Enforcement in the Presence of Tax Avoidance." at *Workshop for the contributors to the Routledge Companion on Tax Avoidance Research*, London, UK, 2015
4. D. Gamannossi degl'Innocenti, "Enforcement of Tax Laws: an Auction Theory Approach." at *XXIV Conferenza Società italiana di economia pubblica*, Pavia, Italy, 2012

Abstract

State intervention emerged in the economy as an instrument to deal with a variety of market failures and led to the development of a wide set of activities: provision of public goods and services, redistribution of wealth across citizens by means of taxes and transfers, education finance, unemployment and health insurance, labour and market regulations.
From one hand, these interventions may cause efficiency losses by distorting agents' choices. From the other hand, efficiency gains may arise from the reduction of the detrimental effects determined by market imperfections. However, the scientific literature also highlights that public intervention in the economy –necessarily made through institutions shaped by the conflicting interests of voters, politicians and bureaucrats– always entails some degree of inefficiency. Hence, whether state intervention enhances efficiency or not, hinges on the particular market imperfection and on the particular implementation of the public intervention considered. The present thesis, acknowledging the potential usefulness of economic policies made by the public sector, is focused on two activities crucial for any state: the collection of revenue, necessary to its very existence and to any policy, and the provision of goods and services made throughout its local branches, with the aim of evaluating its efficiency. Two theoretical contributions are provided that investigate the impact of avoidance and evasion on the self-declaration of liabilities made by taxpayers and on optimal enforcement of tax laws. The third contribution is an empirical inquiry using conditional frontier models to assess the performances of major Italian municipalities.

Chapter 1

Introduction

Revenue collection is affected by an informational problem since liabilities are private information of taxpayers. In developed countries, collection schemes are based on the self-declaration and, in order to contrast the incentive to underreporting, states perform audits and levy fines whenever misleading declarations are detected. In the first contribution of economics to the analysis of tax non-compliance –the seminal article of Allingham and Sandmo (1972)– the evasion decision is analyzed in a portfolio-choice framework. Their work led to a flourishing microeconomics literature that enriched the understanding of the different variables impacting taxpayers' declaration of their liabilities: the psychological and sociological factors affecting agents' behaviour, the influence of legislation and enforcement and how the presence of a variety of non-compliance activities alters the decision. In the thesis two theoretical contributions are presented, both belonging to the latter strand of literature and contemplating avoidance and evasion as non-compliance activity. In the first one, the modelling of taxpayers' joint decision to avoid and/or evade is based on two key assumptions: (i) the decision upon the amounts avoided and/or evaded is performed sequentially and (ii) the avoidance decision is made first. The first feature – that complex decisions are routinely broken down into smaller ones – is often termed *narrow bracketing* (Rabin and Weizsäcker 2009) in the behavioral literature. The second feature – the choice of how to stage the sub-decisions within the larger composite decision – is sometimes termed *decision staging* (Johnson et al. 2012). Crucial with respect to the latter, is whether the taxpayer is more likely to make the avoidance or evasion choice first. This question is thought to depend heavily on mentally focal qualitative features of the choice set (Kahneman 2003). We argue that a focal feature of the choice set is that avoidance is ostensibly legal whereas evasion is illegal, hence, the latter one is not considered until all gainful avenues for the former one have been exhausted. In this setting, the so-called Yitzhaki puzzle identified in the scientific literature – that an increase in the tax rate decreases the level of evaded income and of evaded tax – is proven to hold also in the case of avoidance. However, two major departure from the results obtained previously by the scientific literature arise. First, for a small enough audit probability, evasion is an increasing function of the audit probability. Although tax avoidance is always decreasing in the probability of audit, in some circumstances even the total amount of income lost to evasion and avoidance can be increasing in the probability of audit. Second, holding constant the expected return to evasion, it is not always the case that combined loss of reported income due to avoidance and evasion can be stemmed by increasing the fine rate and decreasing the audit probability. In the second one, the tax agency ensures perfect compliance by enforcing optimal truthtelling probability while the agents choose simultaneously how much to declare, avoid, and evade. Thus, the optimal (counterfactual) level and composition of non-compliance, along with the impact of the relevant parameters, may be inferred from the analysis of the marginal returns of the different concealing options.

In the last part of the thesis it is provided an assessment of productivity efficiency for the local government of the major Italian municipalities. Leveraging on a state-of-the-art methodology, conditional non parametric frontier models (Daraio and Simar 2005 and Daraio and Simar 2007b), the analysis allows to assess the impact of environmental factors on performance and to account for the diversity of exogenous conditions faced by the units. In order to provide a thorough characterization of the productive activity of municipalities, a rich dataset has been gathered from a variety of sources (ISTAT, Ministry of Internal Affairs, Legambiente). Our findings show that northern and central municipalities tend to have better performance with respect to the ones belonging to south and islands. While the investigation does not provide evidence of an effect of debt stock on efficiency, a negative and significant impact of crime is identified. Finally, municipalities characterized by an higher altitude level and/or variability tend to be less efficient.

Notwithstanding major differences in the methodologies employed in the last chapter relative to the first two, a link between the tax compliance and local government efficiency can be traced in the economic literature.

While the framework of portfolio choice allows to understand key insights related to compliance, part of the scientific literature has highlighted the importance of intrinsic willingness to pay taxes, the so-called tax morale, in the determination of non-compliance behaviour (among the others: Schwartz and Orleans 1967; Scholz and Lubell 1998; Alm et al. 1992). This intrinsic motivation has been given both psychological and sociological explanation in terms of the implicit relational contract between taxpayers and government (see e.g., Akerlof 1982 and Kirchler et al. 2008) and found a formalization as a preference for honesty in Myles and Naylor (1996) or as a preference for conformity in Traxler (2010). An extensive experimental evidence (see Torgler 2002 and the references therein) along with field and survey studies (Alm and Torgler 2006; Torgler and Schneider 2009 and Cummings et al. 2009) highlights the direct relation between tax morale and tax compliance and Halla (2012) provided first evidences of a causal link between the two. Furthermore, the idea that price incentives may crowd out intrinsic motivation has been given strong theoretical grounds and is supported by robust empirical evidence (for a theoretical reference see Benabou and Tirole 2003, for an empirical investigation see Frey and Oberholzer-Gee 1997, closely related to the tax compliance setting is Feld and Frey 2007) and puts into perspective the role of enforcement with respect to tax compliance. Quoting Alm et al. (1995) (p.15):

> *[. . .] A government compliance strategy based only on detection and punishment may well be a reasonable starting point but not a good ending point. Instead what is needed is a multi-faceted approach that emphasizes enforcement, as well as such things as positive rewards from greater tax compliance, the wise use of taxpayer dollars, and the social obligation of paying one's taxes.*

Indeed, Hanousek and Palda (2004), Frey and Torgler (2007), Cummings et al. (2009) and Barone and Mocetti (2011) investigated the role of institutions on tax morale and showed a direct relation between the latter and public spending efficiency. Hence, in the literature it has been reported evidence of a positive relation between government performance and tax morale and between tax morale and compliance. While this indirect evidence of a link between government performance and compliance is far from being conclusive, it would be of interest to investigate this relation and the derivation of a reliable measure of government efficiency may represent a first step in this direction. Furthermore, following the literature, it seems reasonable that a more comprehensive compliance strategy should combine traditional enforcement instruments (audit and fines) with compliance enhancing policies aiming to foster honest declaration of liabilities by taxpayers. On this matter, starting from Buchanan (1967, 1976), it has been pointed out the importance of "fiscal exchange", i.e., the need for the government to provide public services suitable in quality and quantity relative to the tax price imposed. Relevant to the present setting, a rich experimental literature found evidence of a link between compliance and public goods provision (Becker et al. 1987; Alm et al. 1992; Alm et al. 1993; Hsu 2008). Indeed, Hsu (2008) identified a positive relationship between compliance and the per-dollar amount of public goods provided. Hence, it seems reasonable that voluntary compliance may be fostered by an improvement of government productive efficiency that would rely, as a crucial step, on the assessment of performances. In that direction some interesting insights are provided in chapter 4.

Chapter 2

Income Tax Avoidance and Evasion: A Narrow Bracketing Approach

2.1 Abstract

We characterize optimal individual tax evasion and avoidance when taxpayers "narrow bracket" the joint avoidance/evasion decision by exhausting all gainful methods for legal avoidance before choosing whether or not also to evade illegally. We find that (i) evasion is an increasing function of the audit probability when the latter is low enough, yet tax avoidance is always decreasing in the probability of audit; (ii) an analogous finding to the so-called Yitzhaki puzzle for evasion also holds for tax avoidance – an increase in the tax rate decreases the level of avoided income and the level of avoided tax; and (iii) that, holding constant the expected return to evasion, it is not always the case that the combined loss of reported income due to avoidance and evasion can be stemmed

by increasing the fine rate and decreasing the audit probability.

2.2 Introduction

Individuals take a variety of actions to reduce their tax liabilities. The UK tax authority, for instance, distinguishes three distinct types of action (HM Treasury & HMRC 2011): those that breach tax law (tax evasion); those that "use the tax law to get a tax advantage that Parliament never intended" (tax avoidance) (p. 3); and those that "use tax allowances for the purposes intended by Parliament" (tax planning) (p. 7). By these definitions, both tax evasion and tax avoidance are responsible for significant losses in public revenue: estimates provided by the UK tax authority put the value of tax avoidance at £2.7 bn. and the value of tax evasion at £4.4 bn. (HMRC 2015). Given the first order significance of tax avoidance, it is of note that the first economic studies relating to tax compliance (e.g., Allingham and Sandmo 1972; Yitzhaki 1974; Christiansen 1980) neglect the possibility of tax avoidance altogether, and the economic literature that followed has largely retained this bias.

In this chapter we introduce tax avoidance into the portfolio model of tax evasion (Yitzhaki 1974). To model the joint tax avoidance/evasion decision we build on insights developed in psychology and behavioral economics. In particular, we allow for a pervasive propensity among human decision makers facing multiple-dimension problems – that of *narrow bracketing*. In our context, a decision maker who narrow brackets

Acknowledgements: We are grateful to participants at the TARC 4th Annual Workshop (Exeter) for helpful comments. This chapter was written while Gamannossi degl'Innocenti was a visitor at Brunel University London, whose hospitality he thanks. Gamannossi degl'Innocenti gratefully acknowledges financial support from the Ministero dell'Istruzione, dell'Università e della Ricerca (cycle XXVIII) and from the European Commission (Erasmus mobility grant 2015-1-IT02-KA103-013713/5).

would decompose sequentially the joint decision {avoidance, evasion} into narrow brackets, e.g., {avoidance} followed by {evasion}. A key feature of narrow bracketing is that the decision maker tends to choose an option in each stage without full regard to the other decisions and circumstances that he or she faces (Rabin and Weizsäcker 2009). Important in this context is whether the taxpayer is more likely to make the avoidance or evasion choice first. This question–as to the order in which a complex decision is mentally staged–is thought to depend heavily on mentally focal qualitative features of the choice set. We argue that a focal feature of the choice set is that avoidance is ostensibly legal whereas evasion is illegal. Indeed, judiciaries have long upheld the right of a citizen to challenge the proper interpretation of tax law and to pay only the tax they owe in law. Thus, while tax avoidance can be seen as the rightful exercise of a basic right by some lights, tax evasion lacks an equivalent interpretation. Accordingly, we suppose taxpayers focus on exhausting opportunities for legal tax avoidance before subsequently focusing on opportunities for illegal evasion.

We are by no means the first to propose that taxpayers distinguish qualitatively between legal and illegal actions, however. This distinction has previously been represented by supposing that a cost owing to social stigma and/or personal guilt is attached to the illegal act of tax evasion. This cost can be financial (e.g., Lee 2001) or psychic (e.g., Gordon 1989). Narrow bracketing offers an alternative perspective: in our model, illegal evasion is not considered until all gainful avenues for legal avoidance have been exhausted.[1] In this way, our chapter relates to a literature

[1]Strictly speaking, the stigma cost approach and ours are not mutually exclusive. To present our approach in the simplest possible light, however, we do not allow for a stigma

on two-stage decision-making (e.g., Blackorby et al. 1970). Unlike this literature, however, we do not seek a sequential approach to the joint avoidance/evasion decision that coincides with the outcome that would obtain were the joint decision assumed to be made simultaneously.

In addition to the legal distinction between avoidance and evasion, we further assume that avoidance is costly whereas evasion is costless. Devising avoidance schemes that reduce a tax liability without ostensibly violating tax law invariably requires a detailed understanding of tax law, coupled with a degree of ingenuity. A classical form of avoidance scheme, for instance, involves the implementation of a circular sequence of self-cancelling option agreements that return the seller to his or her original position, but in the process create an allowable loss. As, however, few taxpayers are equipped to conceive of and implement independently such avoidance schemes, it is necessary to purchase them.[2] Satisfying this demand for tax avoidance is a substantial industry dedicated to the development and marketing of avoidance schemes (see, e.g., Sikka 2012; Committee of Public Accounts 2013). By contrast, many forms of tax evasion require no technical or legal expertise. Intentionally understating income on the tax return, for instance, may readily be performed independently. We find that, in two respects, allowing for tax avoidance importantly changes the characteristics of optimal tax evasion. First, under plausible conditions, evasion is an *increasing* function of the probability of audit. Second, we re-examine the finding of (Christiansen, 1980, p. 391) that "if the fine is increased, but the efforts to detect tax evaders are

cost.

[2]People not only have difficulties in understanding tax law but also show poor knowledge about tax rates and basic concepts of taxation. For evidence that people often significantly underestimate marginal tax rates, see, for example, Gideon (2015).

adjusted so as to keep the expected gain from tax evasion unaltered, risk averters will always reduce their tax evasion." We are again able to prove this result, yet in our model it is the total amount of lost tax (through both avoidance and evasion) that is economically pertinent. When we consider both avoidance and evasion, it is possible that taxpayers declare more income if the probability of audit is increased, and the fine decreased, holding the expected gain from tax evasion constant. This chapter adds to the small, but growing, economic literature on tax avoidance. The two closest analyses to ours are Alm and McCallin (1990) and Alm (1988a). The former describes avoidance and evasion as risky assets – each asset has a return characterized by a mean and variance, and the interdependence between the two returns is characterized by a covariance – while the latter characterizes avoidance as a riskless, albeit costly, asset. Whereas both of these analyses consider the simultaneous determination of evasion and avoidance, in our framework, we argue that these are chosen sequentially. Different from Alm and McCallin, we model the mean, variance, and covariance of evasion and avoidance, rather than taking these quantities as exogenous. Unlike in Alm (1988a), we take avoidance to be risky, owing to the possibility of effective anti-avoidance measures by the tax authority. Much of the remaining literature on tax avoidance is, however, concerned with whether income tax has "real" effects upon labor supply, or simply leads to changes in the "form" of compensation (e.g., Slemrod and Kopczuk 2002; Piketty et al. (2014); Slemrod (1995)). Accordingly, in these studies the term "tax avoidance" typically refers to all form-changing actions that reduce a tax liability.[3] This definition

[3] For a detailed discussion of these "form-changing" actions see, e.g., Stiglitz (1986) and Slemrod and Yitzhaki (2002).

overlaps with ours, but is broader in the sense that it also includes actions that fall in to our notion of tax planning. By this broader definition Lang et al. (1997) estimate that tax avoidance costs the German exchequer an amount equal to around 34 percent of income taxes paid. The plan of the chapter is as follows: in section 2.3 we motivate the key behavioral assumptions behind our analysis, from which section 2.4 develops a formal model. Section 2.5 performs the main analysis and 2.6 compares our findings to the literature. We extend the model in section 2.7 to allow for risk in the tax authority's efforts to illegalize avoidance schemes, and section 2.8 concludes. All proofs are in the Appendix.

2.3 Deciding to Avoid and/or Evade Tax

Two key features of our modelling of the joint decision to avoid and/or evade are that (i) the taxpayer makes the avoidance and evasion decisions sequentially; and (ii) that the avoidance decision is made first. The first feature – that complex decisions are routinely broken down into smaller ones – is often termed *narrow bracketing* in the behavioral literature. The second feature – the choice of how to stage the sub-decisions within the larger composite decision – is sometimes termed *decision staging* (Johnson et al. 2012). We discuss each of these features in turn.

2.3.1 Narrow Bracketing

A mass of evidence suggests that people narrowly bracket: a decision maker who faces a multi-dimensional decision tends to break the decision down sequentially, proceeding at each stage to isolate a single dimension of the problem without full regard to the other dimensions of the problem. In the context of monetary risk, Tversky and Kahneman (1981) present an experiment that demonstrates how powerful this propensity is. In their experiment, people narrowly bracket even when faced with only a pair of independent simple binary decisions that are presented on the same sheet of paper. Narrow bracketing lies at the heart of current explanations of phenomena such as the stock market participation puzzle (Barberis et al. 2006), the equity premium puzzle (Benartzi and Thaler 1995; Gneezy and Potters 1997) and choice among lotteries (Battalio et al. 1990; Langer and Weber 2001).

Cognitive limitations – in perception, attention, memory, and analytical processing – are thought to be an important reason for narrow bracketing (Read et al. 1999). Accordingly, in the experiment of Read et al. (2001) subjects who were required to resolve a complex choice problem sequentially actually made better choices than those subjects who were required to proceed simultaneously. Clearly, however, decision making outcomes under narrow bracketing can, in other contexts, appear worse than those arrived at from a wide bracketing perspective. For instance, decision making under narrow bracketing may violate first-order stochastic dominance (Rabin and Weizsäcker 2009) and, in a famous example, the "one day at a time" bracketing observed among New York cab drivers fails to maximize earnings per hour across days (Camerer et al. 1997). Thus, while the desirability of narrow bracketing is still the subject of academic debate (see, e.g., Kőszegi and Rabin 2009), the pervasiveness of the phenomenon is not in doubt. It is thus of relevance to understand the nature of the joint avoidance and evasion decision under narrow bracketing.

2.3.2 Decision Staging

Having established that taxpayers may well mentally separate the joint avoidance and evasion decision there remains the question as to how this is done. According to Kahneman (2003), the way individuals will choose to stage or frame a decision is heavily shaped by the features of the situation at hand that come to mind most easily – to use the technical term, by the features that are most "accessible." This notion is supported in the context of tax-related decision making by McCaffery and Baron (2004). These authors employ a slightly different terminology – the *isolation effect* – which, however, refers to the tendency of respondents in their study to "decide complex matters by responding to the most salient or obvious aspect of a choice set or decision problem."In the context of the joint avoidance and evasion decision, we argue that the most accessible feature of the choice set is that tax avoidance and evasion are qualitatively distinct: one is legal, and the right to practice it has often been defended by the judiciary, whereas the other is a crime. Courts on both sides of the Atlantic have for many years upheld the right of citizens to challenge the interpretation of tax law (Barker 2009; Prebble and Prebble 2010). In 1936 Lord Tomlin surmised that "[e]very man is entitled, if he can, to order his affairs so that the tax attaching under the appropriate Acts is less than it otherwise would be. If he succeeds in ordering them so as to secure this result, then, however unappreciative the Commissioners of Inland Revenue or his fellow taxpayers may be of his ingenuity, he cannot be compelled to pay an increased tax." Similarly, in the United States, Judge Learned Hand stated in 1947 (in Commissioner vs. Newman) that "[t]here is nothing sinister in so arranging one's affairs as to keep taxes as low as possible. Everybody does so, rich or poor; and all do right, for nobody owes any public duty to pay more than the law demands."There is evidence that these traditional legal arguments continue to affect public sentiment towards tax avoidance. In the qualitative study of Kirchler et al. (2003), participants relate tax avoidance to law- ful acts enabling tax reduction, to cleverness, and to costs. Tax evasion, by contrast, is associated with illegal acts such as fraud, criminal prosecution, risk, tax audits, punishment, penalty, and the risk of detection. In this sense, we argue that, for many taxpayers, tax avoidance is qualitatively preferred to evasion. Accordingly, we argue that taxpayers would exhaust the scope for legal avoidance before subsequently deciding whether or not they addition- ally wish to evade illegally (rather than the other way round).

2.4 Model

A taxpayer has an income (wealth) w and faces a tax on income given by tw, where $t \in (0, 1)$. Taxpayers behave as if they maximize expected utility, where utility is denoted by $U(z) = \log z$.[4] The taxpayer's true income is not observed by the tax authority, but the taxpayer must declare an amount $x \in [0, w]$. The taxpayer can choose to avoid paying tax on an amount of income $A \in [0, w]$, and subsequently to evade illegally an amount of income $E \in [0, w - A]$, so $x = w - E - A$. Evasion is financially costless but avoidance technology must be bought in a market in which

[4]Thus taxpayers are risk averse and have a constant (unit) co-efficient of relative risk aversion. We adopt the logarithmic form for reasons of analytic tractability, though we note that the assumption of constant relative risk aversion commands considerable empirical support (see, e.g., Chiappori and Paiella 2011.

"promoters" sell avoidance schemes to "users".[5] A common feature of this market is the "no saving, no fee" arrangement under which the price received by a promoter is linked to the amount by which their scheme stands to reduce the user's tax liability. Although sys- tematic information regarding the contractual terms of avoidance schemes is scarce, we understand from a detailed investigation in the UK that, for the majority of mass-marketed schemes, the fee is related to the reduction in the annual theoretical tax liability of the user, not the *expost* realization of the tax saved (Committee of Public Accounts 2013). Thus, the monetary risks associated with the possible subsequent detection and termination of a tax avoidance scheme are borne by the user.[6] Accordingly, we assume that the promoter's fee is a proportion $\phi \in (0, 1)$ of the amount by which the taxpayer's tax liability stands to be reduced, tA. In this way, ϕ may be interpreted as measuring the degree of competition in the market for tax avoidance schemes, with lower values of ϕ indicating the presence of stronger competitive forces. Although traditional arguments around the morality of tax avoidance continue to affect importantly public sentiment, there has nonetheless been a discernible shift in the attitudes of the judiciary, beginning in the 1980s (Stevens 2013). Increasingly, courts apply a purposive interpretation, as summarized by Judge Ribeiro, who states (in *Collector of Stamp Revenue vs. Arrowtown Assets Ltd.*) who that "the ultimate question is whether the relevant statutory provisions, construed purposively, were intended to apply to the transaction, viewed

[5]For analyses of the market for tax advice see, e.g., Reinganum and Wilde (1991) and Damjanovic and Ulph (2010).

[6]It is apparent that such arrangements give promoters incentives to misrepresent the level of risk involved in particular schemes. Consistent with this point, (Committee of Public Accounts, 2013, p.11) indeed finds evidence of such mis-selling.

realistically." Armed with this purposive interpretation of the law, tax authorities now routinely seek to have avoid ance schemes ruled illegal. Yet taxpayers may legitimately continue to use an avoidance scheme while the (often lengthy) process of shutting it down is ongoing. Moreover, if the scheme is eventually declared illegal, the tax authority can only seek the amount of tax that was properly due (it cannot levy fines retrospectively). Inherent in our definition of tax avoidance (as distinct from tax planning) is that–should the tax authority learn of the scheme–it will consider it illegal. The taxpayer's income declaration is audited with probability $p \in (0,1)$. If audited, E and A are observed and the taxpayer has to pay $[1+f]\,tE$ on account of the amount of evaded tax, where $f > 0$ is the fine rate. The tax authority mounts a legal challenge to the avoidance scheme, which is successful with probability p_L. In the event that the legal challenge is successful, the tax authority obtains the right to reclaim the tax owed (but cannot levy a fine). In this case, instead of paying tx in tax, the taxpayer must instead pay $t\,[x+A]$.

The taxpayer's expected utility is therefore given by

$$\mathbf{E}U(A,E) = [1-p]\,U(w^n) + pp_L U(w^{a_s}) + p\,[1-p_L]\,U(w^{a_u}), \qquad (2.1)$$

where w^n is the taxpayer's wealth in the state in which they are not audited, w^{a_s} is the taxpayer's wealth in the state in which they are audited and the tax authority's legal challenge is successful, and w^{a_u} is the taxpayer's wealth in the state in which they are audited and the tax authority's legal challenge is unsuccessful:

$$
\begin{aligned}
w^{a_s}(A,E) &= w - t[w-E] - [1+f]\,tE - \phi tA; &(2.2)\\
w^{a_u}(A,E) &= w - t[w-A-E] - [1+f]\,tE - \phi tA; &(2.3)\\
w^{n}(A,E) &= w - t[w-A-E] - \phi tA. &(2.4)
\end{aligned}
$$

A key distinguishing factor between evasion and avoidance in this context is that avoidance entails a cost ϕtA in all states of the world. Thus, if avoidance is detected and the scheme closed down a taxpayer is worse-off for having chosen to avoid, even though they are not fined on avoided income. To ensure that the amount of taxes, fines and fees never exceeds a taxpayer's wealth for any $A + E \in [0, w]$ we must assume $[1-t]/t > \max\{\phi, f\}$.

We suppose that taxpayers choose their preferred level of avoidance and evasion sequentially: gainful opportunities for tax avoidance are exhausted before the taxpayer decides whether to engage additionally in evasion. Thus, taxpayers first choose avoided income as

$$A^* = \arg\max_A \mathbf{E}U(A, 0)\,; \tag{2.5}$$

and then evaded income as

$$E^* = \arg\max_E \mathbf{E}U(A^*, E)\,. \tag{2.6}$$

2.5 Analysis

We now present an analysis of the model of the previous section. For analytic tractability, we shall consider the special case of the model with

$p_L = 1$, such that legal challenges by the tax authority are always successful. In a later section we shall demonstrate numerically how the results with $p_L = 1$ relate to the results for the more general case with $p_L < 1$.

To begin, it is helpful to define the function $R(z) = [1 - z]/z$, such that, e.g., $R(p)$ is the classical odds ratio found in decision theory. We may then state our first Proposition:

Proposition 1 *An interior optimum for avoidance and evasion satisfies*

$$A^* = \frac{pR(t)}{1-\phi}\left[R(p)\,R(\phi) - 1\right]w$$

and

$$E^* = \frac{pR(t)}{1-\phi}\frac{[1-p]\,[1-fR(\phi)]}{f}w$$

where

$$R(p)R(\phi) > 1 > fR(\phi);$$
$$\frac{pR(t)}{1-\phi}\frac{[1-p][1-fR(\phi)] + f[R(p)R(\phi)-1]}{f} < 1.$$

Proposition 6 gives closed-form expressions for optimal avoidance and evasion when both are at an interior maximum, and the conditions needed for a such an interior maximum to arise. The first sequence of inequalities at the bottom of the Proposition guarantee that $A^*, E^* > 0$. The left-side inequality, $R(p)\,R(\phi) > 1$, is the condition that the avoidance gamble be better than fair. As a necessary condition for both inequalities to hold it must be that $R(p)\,R(\phi) > fR(\phi)$, which implies $R(p) > f$. This is the standard restriction in the portfolio model of tax evasion that the evasion gamble be better than fair. The right side inequality at the bottom of the Proposition ensures that $A^* + E^* < w$. To gain insight into how A^* and E^* are related, note that we may write one as a (linear) function of the other:

$$E^{*}\left(A^{*}\right)=\frac{p\left[wR\left(t\right)-\phi A^{*}\right]\left[R(p)-f\right]}{f}-pA^{*}. \tag{2.7}$$

From (2.7) we note that

$$\frac{\partial E^{*}}{\partial A^{*}}=-\frac{fR(\phi)+R(p)}{f\left[1+R(p)\right]\left[1+R(\phi)\right]}<0, \tag{2.8}$$

so the amount of evaded income E^{*} is negatively related with the amount of avoided income A^{*}. This finding matches that of Alm et al. (1990) using data from Jamaica, who find that evasion and avoidance are substitutes. We now consider the comparative statics of optimal avoidance and evasion:

Proposition 2 *At an interior optimum for avoidance and evasion it holds for A^{*} that*

$$\begin{aligned}
\frac{\partial A^{*}}{\partial w} &= \frac{A^{*}}{w}>0;\\
\frac{\partial A^{*}}{\partial t} &= -\frac{A^{*}}{t[1-t]}<0;\\
\frac{\partial A^{*}}{\partial f} &= 0;\\
\frac{\partial A^{*}}{\partial \phi} &= -\frac{wR(t)\left[p+(1-p)R(\phi)^{2}\right]\left[1+R(\phi)\right]^{2}}{R(\phi)^{2}}<0;\\
\frac{\partial A^{*}}{\partial p} &= -\frac{R\left(t\right)w}{\phi\left[1-\phi\right]}<0;
\end{aligned}$$

and for E^* that

$$
\begin{aligned}
\frac{\partial E^*}{\partial w} &= \frac{E^*}{w} > 0; \\
\frac{\partial E^*}{\partial t} &= -\frac{E^*}{t\,[1-t]} < 0; \\
\frac{\partial E^*}{\partial f} &= -\frac{E^*}{[1-fR(\phi)]\,f} < 0; \\
\frac{\partial E^*}{\partial \phi} &= \frac{f\,[R(\phi)]^2+1}{[1-\phi]\,[1-fR(\phi)]}E^* > 0; \\
\frac{\partial E^*}{\partial p} &= \frac{R(p)-1}{1-p}E^* \gtreqless 0 \Longleftrightarrow R(p) \gtreqless 1.
\end{aligned}
$$

Proposition 2 is derived via straightforward differentiation of the expressions for A^* and E^* in Proposition 6 so we omit the proof. Beginning with the comparative statics of A^*, we see that wealthier people are predicted to avoid more income than less wealthy people. The second result is an extension of the well-known Yitzhaki paradox for evasion to the case of avoidance – avoided income falls as the tax rate is increased. The intuition for this result is analogous to that for evasion: a higher marginal tax rate makes the taxpayer feel poorer, and thereby more risk averse. An increase in the competitiveness of the market for avoidance schemes (a decrease in ϕ) increases avoided income, and an increase in the probability of audit decreases avoided income. Of course, knowing $\partial A^*/\partial t < 0$ does not warrant that the total tax avoided, tA, also falls. It is straightforward to show, however, that

$$
\frac{\partial tA^*}{\partial t} = -\left\{\frac{1}{t\,[1-t]} - 1\right\} A^* < 0. \tag{2.9}
$$

Turning to evasion, the logic of the chain rule implies that, for an arbitrary exogenous variable z, it must hold that

$$\frac{\partial E^*}{\partial z} = \left.\frac{\partial E^*}{\partial z}\right|_{A^*=cons.} + \frac{\partial E^*}{\partial A^*}\frac{\partial A^*}{\partial z}, \tag{2.10}$$

where the first term on the right side is the direct effect of z on evasion, and the second term captures the indirect effect on evasion arising from the effects of z upon avoidance. Intuitively, the indirect effect is the income effect imparted upon the evasion choice by movements in avoidance. Noting that $\partial E^*/\partial A^* < 0$ it follows that if $\partial A^*/\partial z$ and $\partial E^*/\partial z|_{A^*=cons.}$ are of the same sign – as turns out to be the case for each of the variables $\{w, t, f, p\}$ – then the direct and indirect effects in equation (2.10) oppose each other. This observation notwithstanding, the first three results in Proposition 2 – those for $\{w, t, f\}$ – are each unambiguous and consistent with Yitzhaki (1974). Unambiguity in this context arises as the direct effect can be shown to always dominate the indirect effect. To take the tax rate as an example, we find the direct effect – using (2.7) – as

$$\left.\frac{\partial E^*}{\partial t}\right|_{A^*=cons.} = -\frac{w\,[1+R(t)]^2\,[R(p)-f]}{f\,[1+R(p)]} < 0, \tag{2.11}$$

Combining equation (2.11) with $\partial A^*/\partial t$ in Proposition 6 and $\partial E^*/\partial A^*$ we can rewrite the direct effect in terms of the indirect effect,

$$\left.\frac{\partial E^*}{\partial t}\right|_{A^*=cons.} = -\frac{[R(p)-f]\,[1+R(p)]\,R(\phi)}{[R(p)+fR(\phi)]\,[R(p)R(\phi)-1]}\frac{\partial E^*}{\partial A^*}\frac{\partial A^*}{\partial t}, \tag{2.12}$$

such that, by equation (2.10), we obtain an alternative form for $\partial E^*/\partial t$ to that given in Proposition 2:

$$
\begin{aligned}
\frac{\partial E^*}{\partial t} &= -\frac{[R(p)-f][1+R(p)]R(\phi)}{[R(p)+fR(\phi)][R(p)R(\phi)-1]}\frac{\partial E^*}{\partial A^*}\frac{\partial A^*}{\partial t}+\frac{\partial E^*}{\partial A^*}\frac{\partial A^*}{\partial t} \\
&= -\frac{R(p)[1+R(\phi)][1-fR(\phi)]}{[R(p)+fR(\phi)][R(p)R(\phi)-1]}\frac{\partial E^*}{\partial A^*}\frac{\partial A^*}{\partial t}<0
\end{aligned}
\tag{2.13}
$$

As well as evaded income being decreasing in the tax rate, it is straightforward to show that evaded tax, tE^*, is decreasing in the tax rate too. As it has no direct effect on evasion, the effect of competition in the market for avoidance (as captured by ϕ) is given by (2.10) as simply $\partial E^*/\partial\phi = [\partial E^*/\partial A^*]\,[\partial A^*/\partial\phi] > 0$. Thus, a decrease in the competitiveness of the avoidance industry increases evasion. The final finding is that tax evasion is increasing in the probability of audit if $R\,(p) > 1$ (equivalently, $p < 0.5$) and decreasing otherwise. In this case there are again competing direct and indirect effects upon evasion, but now the direct effect does not always dominate the indirect effect. Following the same steps as in the tax rate example above, we obtain

$$
\frac{\partial E^*}{\partial p} = \frac{[R(p)-1]\,[1-fR(\phi)]}{R(p)+fR(\phi)}\frac{\partial E^*}{\partial A^*}\frac{\partial A^*}{\partial p}. \tag{2.14}
$$

From (2.14) it is immediate that the direct effect dominates when $R(p) < 1$ and the indirect dominates when $R(p) > 1$.

How plausible is the condition $p < 0.5$ required for evasion to be increasing in the probability of audit? A priori it appears highly plausible given that the IRS audits only around 0.96 percent of individual tax returns filed in calendar year 2012 were examined (Internal Revenue Service 2014). If audits are concentrated on the 20 percent or so of people in the United States who are self-employed, the probability for this group would rise to 4.8 percent, still well below the 50 percent level. We now wish to characterize the total level of undeclared income, $A^* + E^*$:

Proposition 3 *An interior optimum for avoidance and evasion it holds for* $A^*+ E^*$ *that*

$$\begin{aligned}
\frac{\partial\left[A^*+E^*\right]}{\partial w} &= \frac{A^*+E^*}{w}>0; \\
\frac{\partial\left[A^*+E^*\right]}{\partial t} &= -\frac{A^*+E^*}{t[1-t]}<0; \\
\frac{\partial\left[A^*+E^*\right]}{\partial f} &= \frac{\partial E^*}{\partial f}<0; \\
\frac{\partial\left[A^*+E^*\right]}{\partial \phi} &= \frac{p^2 w R(t)\left\{R(p)\left\{1-f-fR(p)\left[R(\phi)\right]^2\right\}-f\right\}}{\left[1-\phi\right]^2 f} \gtrless 0 \\
&\iff R(p)\left\{1-f-fR(p)\left[R(\phi)\right]^2\right\}-f \gtrless 0; \\
\frac{\partial\left[A^*+E^*\right]}{\partial p} &= \frac{wR(t)\left[1+R(\phi)\right]\left\{R(p)\left[1-f-2fR(\phi)\right]-1-f\right\}}{f\left[1+R(p)\right]R(\phi)} \gtrless 0 \\
&\iff R(p)\left[1-f-2fR(\phi)\right]-\left[1+f\right] \gtrless 0.
\end{aligned}$$

The first three results of Proposition 3 follow immediately from Proposition 2, for the comparative static effects for both A^* and E^* go in the same direction. Both of the remaining two effects – those for ϕ and p – may go in both directions, however. Unlike the condition for E^* to increase in p, though, the condition needed for $w - x$ to increase in p seems far from being satisfied empirically. In particular, it requires setting an especially low f, which, in turn, forces the tax rate to be implausibly high. A similar remark applies to the condition needed for $w - x$ to increase in ϕ. The results of Proposition 3 allow us to characterize readily the comparative statics of declared income x^*). For all exogenous variables except w we obtain that the effect for declared income will take the opposite sign to the effect for total undeclared income, i.e., $\partial x^*/\partial(\cdot) = -\partial\left[w - x^*\right]/\partial(\cdot)$. For w, however, we obtain $\partial x^*/\partial w = x^*/w > 0$.

As a final perspective on the properties of optimal avoidance and evasion, we may characterize the properties of the proportion (or share) of unreported income that is avoided: $s_A \equiv A/\left[A+E\right]$. This shall be instructive when we come to compare our results with the existing literature.

Proposition 4 *At an interior optimum for avoidance and evasion it holds that*

$$\begin{aligned}
\frac{\partial s_A}{\partial t} &= \frac{\partial s_A}{\partial w} = 0; \\
\frac{\partial s_A}{\partial f} &= \phi p^2 \left[1-p-\phi\right] R(p)\vartheta > 0; \\
\frac{\partial s_A}{\partial \phi} &= -p^2 \left[1-p-fp\right] R(p)\vartheta < 0; \\
\frac{\partial s_A}{\partial p} &= -f\left[\phi p\right]^2 \left[1-fR(\phi)\right]\left[1+R(p)^2R(\phi)\right]\vartheta < 0;
\end{aligned}$$

where $\vartheta \equiv \left\{\frac{A^*}{f[A^*+E^*][1-p-\phi]}\right\}^2 > 0.$

Proposition 4 clarifies that the share of undeclared income that is avoided is independent of the tax rate (t) and the taxpayer's wealth (w). This follows from the observation in Proposition 6 that w and $R\left(t\right)$ enter both avoidance and evasion as multiplicative factors. We find that the probability of audit unambiguously reduces the share of undeclared income that is avoided, even though the effect of p on evasion can be of either sign. The results for the effects of the fine rate and the cost of avoidance follow directly from Proposition 2.

2.6 Comparison with the literature

We now compare the findings of the previous section to the existing literature. First, we consider Alm et al. (1990). These authors find that the quantity $[1 - ft]^{-1}$ is negatively related to evaded income, implying that an increase in either f or t reduces evasion (consistent with Proposition 2).They also find, like us, that avoided income is decreasing in the cost of avoidance (as measured in our model by the parameter ϕ). A caveat,however, is that Alm, Bahl, and Murray identify avoidance as a riskless asset, somewhat different from the definition of avoidance as a risky asset we employ here. Second, we may compare our findings for optimal evasion to those of Yitzhaki (1974)'s canonical model of tax evasion. Our findings for the effect of wealth, the tax rate and the fine rate on evasion are consistent with Yitzhaki, but the finding that evasion may increase in the probability of audit is different from that in Yitzhaki (where evasion is always decreasing in the probability of audit). Third, we may compare our findings to those of Alm and McCallin (1990), who report comparative statics results for reported income (x) and for the share of undeclared income that is avoided (s_A). Like these authors, we find that higher fines for evasion increase reported income and increase the share of undeclared income that is avoided. Different from these authors, however, we retain the well-known result of Yitzhaki (1974) that a tax rate rise will increase reported income (whereas Alm and McCallin report the opposite relationship) and, thoughs Alm and McCallin find that a tax rate rise increases the share of undeclared income that is avoided, we find that this share is independent of the tax rate. It can be shown that this independence is robust to allowing for $p_L < 1$, and allowing for a coefficient of constant relative risk aversion that is different from unity. We are unable, however, to follow Alm and McCallin in examining the comparative statics effects of quantities such as the mean and variance of the return to evasion and avoidance, however, as in our model these quantities are determined endogenously. Last, we may compare our findings to those of the theoretical model of Alm (1988a), albeit an important difference between his model and ours is that he models avoidance as a riskless asset. Alm presents comparative statics results for the quantities $w - E$ and the share $s_x \equiv [w - A - E] / [w - E] = x / [w - E]$, i.e., the fraction of income net of evasion that is avoided. In his very general framework, Alm finds all comparative statics for the share s_x to be ambiguous in sign. We similarly find that the effects of ϕ and p on s_x are ambiguous, but we find that $\partial s_x / \partial f > 0$ and $\partial s_x / \partial t > 0$. We find that s_x is independent of a taxpayer's wealth: $\partial s_x / \partial w = 0$. The only other two clear-cut results in Alm (1988a) are that evasion is decreasing in the fine rate and in the audit probability. In our model the first of these effects is preserved, but we find that evasion can be increasing in the probability of audit.

2.6.1 Audit probability vs. fine rate

As a final comparison to the literature, we consider the finding of Christiansen (1980) that, for a constant expected return to evasion, the amount evaded is always reduced by increasing the fine rate and by decreasing the audit probability. Following Christiansen (1980), we first restrict analysis solely to evasion. For a given level of avoidance the expected return to evasion is given by $\mu_E \equiv p\,[1 + f] - 1$. Holding this constant by appropriate variation of f, and differentiating E^* with respect to p, we obtain:

Proposition 5 *An interior optimum for avoidance and evasion it holds that*

$$\left.\frac{\partial E^*}{\partial p}\right|_{\mu_E = cons.} = \frac{wR(t)\left([1 - R(p)]\, f^2 R(\phi) + [1 + 2f]\, R(p) - f\right)}{f^2\,[1 - \phi]\,[1 + R(p)]} > 0.$$

According to Proposition 5 we are able to replicate Christiansen's finding: it always worsens evasion to raise the audit probability and lower the fine rate, holding the expected return to evasion fixed. In the context of a model containing both avoidance and evasion, however, what will be relevant to a tax authority seeking to maximize tax revenue is the effect of varying p and f on the total level of income that does not get taxed. On this question we have that:

$$\left.\frac{\partial \left[A^{*} + E^{*}\right]}{\partial p}\right|_{\mu_E = cons.} = \frac{wR(t)\left\{[1-p]\left\{1 + f\left[3 - 2fR\left(\phi\right)\right]\right\} - f\left[1+f\right]\right\}}{\left[1-\phi\right]f^{2}} \gtrless 0. \tag{2.15}$$

As the right side of equation (2.15) can take either sign, depending upon parameter values, we now no longer find that raising fines is always superior to raising audit probability. Intuitively, this finding stems from the observation that increasing the fine rate only affects the evasion decision, whereas increasing the audit probability affects both avoidance and evasion.

2.7 Probabilistic Anti-Avoidance Outcomes

Up until this point, the analysis has been undertaken with the simplifying assumption that, if the tax authority mounts a legal challenge to the avoidance scheme, its challenge is always successful. While important in securing a tractable model, clearly tax authorities are not always successful in their attempts to shut-down avoidance schemes, so it is of interest to understand how this consideration affects our findings.

Solving for A^{*} using the definition in equation (2.5) and the full expression for expected utility given in (1) we obtain

$$A^* = \frac{pp_L R(t)}{1-\phi}\left[R(pp_L)R(\phi)-1\right]w, \tag{2.16}$$

which can be obtained from the solution for A^* given in Proposition 6 (for the case of $p_L = 1$) simply by replacing p with pp_L.[7] It follows that the comparative statics results for A^* given in Proposition 2 continue to hold, and that the effects of p_L upon avoidance are analogous to those of p. The solution for E^* coming from (2.6) is complex, however. We note, though, that the the taxpayer's wealth and the tax rate both still enter the solution multiplicatively – as they do also for A^* in equation (2.16) – so these two variables stay independent of the share of undeclared income that is avoided, s_A.

To make further progress we assess the properties of optimal evasion via a numerical optimization procedure. Figure 1 depicts optimal avoidance and evasion as p_L is allowed to vary on the unit interval. [8]. For very low values of p_L in the interval denoted $[0, \hat{p}_L|_{E=0}]$ avoidance is seen to be maximal, and the taxpayer does not evade. In a second interval, denoted in the figure by $(\hat{p}_L|_{E=0}, \hat{p}_L|_{A+E=w}]$, the taxpayer both avoids and evades, and reports no income ($x \geq 0$ is binding). In a third interval, denoted $(\hat{p}_L|_{A+E=w}, 1]$, the taxpayer again both avoids and evades, but

[7] Whereas the expression for A^* given in Proposition 6 is the unique solution to a first order condition linear in A^*, the expression for A^* in (2.16) is one of a pair of solutions to a first order condition quadratic in A^*. The other solution to this first order condition is $A^* = -wR(t)/[1-\phi] < 0$, which however, may be dismissed as, by definition, $A^* \geq 0$.

[8] The parameter values that produce Figure 1 are: $w = 10, p = 0.5, t = 0.8, f = 0.1, \phi = 0.22$. We note that these values are chosen purely to illustrate cleanly the full range of possible outcomes of the model. As is well-known, models such as ours, which implicitly assume taxpayers know the true probability of audit, significantly over-predict non-compliance if calibrated realistically (see, e.g., Alm et al. 1992 , footnote 3). This difficulty does not appear especially consequential in this context, however, for insights such as probability weighting (Kahneman and Tversky 1979) have been shown to dramatically reduce predicted levels of non-compliance, while not importantly affecting its comparative static properties (these being our interest in this paper)

now $x > 0$ (this is the case to which our comparative statics analysis applies). Within this interval we observe that optimal evasion increases as the probability of a successful legal challenge increases. This is as expected, for an increase in p_L leaves the returns to evasion unaffected, but reduces the returns to avoidance, making evasion more attractive relative to avoidance.

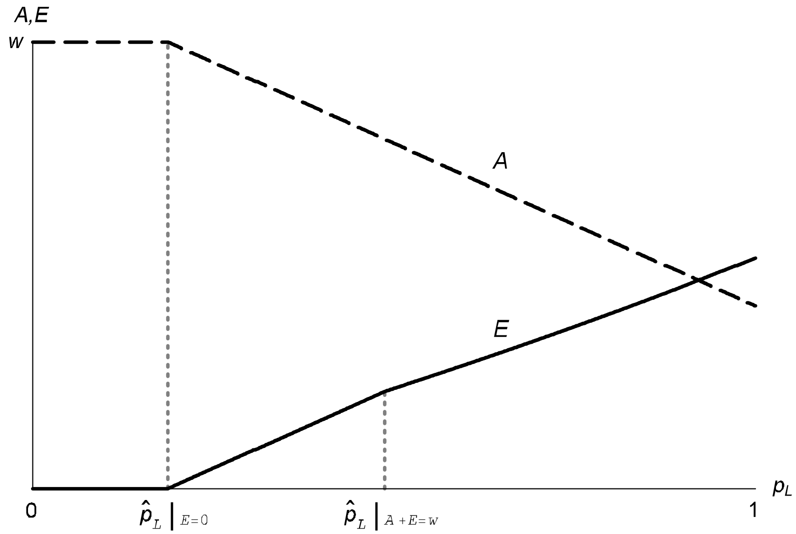


Figure 1: Optimal avoidance and evasion for $p_L \in [0, 1]$.

In Figure 2 we explore the effect of varying p_L on our earlier finding that evasion can be increasing in the probability of audit.[9] On the interval of Figure 2 where both optimal evasion and avoidance are interior, we

[9] The parameter values that produce Figure 2 are: $w = 10, t = 0.85, f = 0.11, \phi = 0.17, p_L \in \{0.7, 1\}$. The same qualitative conclusions obtain if the parameter values used to draw Figure 1 are instead used, but these alternative parameter values yield improved visual clarity.

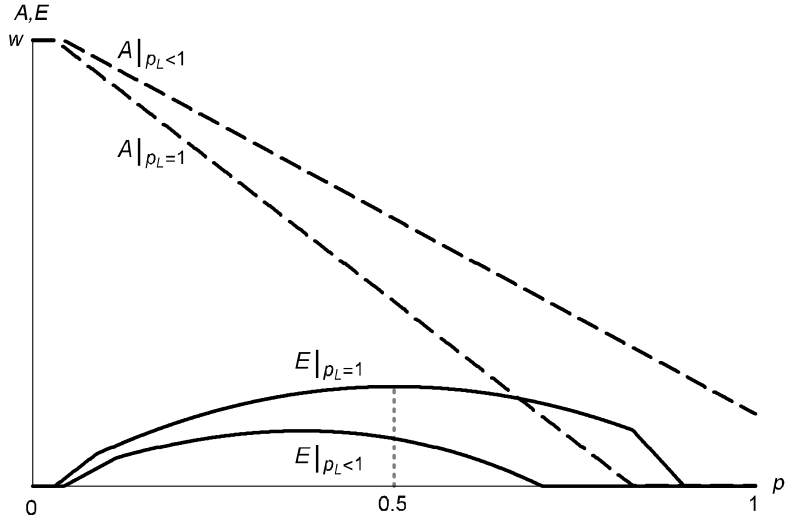


Figure 2: Optimal avoidance and evasion for $p_L < 1$ and $p_L = 1$.

see that reducing p_L below unity reduces to a value below one-half the threshold audit probability above which evasion is decreasing in p. Thus lower values of p_L imply a smaller set of parameter values for which evasion is observed to be increasing in audit probability. Other numerically generated results we have analyzed – which we do not report here for brevity – indicate that the qualitative nature of the results given in propositions 2-4 continue to hold. In particular, the taxpayer's wealth and the tax rate continue to act as multipliers in the expressions for optimal avoidance and evasion, and $w - x$ may be either an increasing or decreasing function in ϕ and p.

2.8 Conclusion

Although the economic literature has largely limited itself to the study of tax evasion, tax avoidance is empirically observed alongside tax evasion. We therefore examine the choice of a taxpayer of how much tax to avoid and how much to evade, under the assumptions that (i) a taxpayer narrow brackets the joint avoidance/evasion decision – breaking the decision down into separate avoidance and evasion sub-decisions, and taking the first of these two sub-decisions in isolation from the second; and (ii) that a taxpayer will decide first on whether and how much tax to avoid legally before deciding whether and how much tax to evade illegally. Among our results are, first, that an analogous finding to the so-called Yitzhaki puzzle for evasion also holds for tax avoidance – an increase in the tax rate decreases the level of avoided income and the level of avoided tax. Second, for a small enough audit probability, evasion is an increasing function of the audit probability. Although tax avoidance is always decreasing in the probability of audit, in some circumstances even the total amount of income lost to evasion and avoidance can be increasing in the probability of audit. Last, holding constant the expected return to evasion, it is not always the case that combined loss of reported income due to avoidance and evasion can be stemmed by increasing the fine rate and decreasing the audit probability.

We finish with some possible avenues for future research. First, it would be of interest to allow for imperfect audit effectiveness, as in Rablen (2014) and Snow and Warren (2005b), for it might be that evasion and avoidance differ in the amount of tax inspector time required to detect them. Second, it would be extremely useful to perform an empirical test on the results derived from the model. Unfortunately, no suitable data is available to undertake such investigation at the moment. However, an alternative to this end would be to set up an experimental study. Indeed, further research along these lines could provide valuable insights to foster our understanding of non-compliance decision making. A last suggestion is to embed the model within a general equilibrium framework (see, e.g.,Alm and Finlay 2013), for the partial equilibrium setting explored here may miss some important wider interactions between avoidance and evasion that should properly be accounted for.

2.9 Appendix

Proof of Proposition 6: From (1) we have that

$$\begin{aligned}\frac{\partial \mathbf{E}U\left(A,0\right)}{\partial A} &= \frac{[1-p][1-\phi]}{A[1-\phi]+wR(t)} - \frac{\phi p}{wR(t)-\phi A} \qquad (17)\\ &= \phi p\left\{\frac{R(p)R(\phi)}{A[1-\phi]+wR(t)} - \frac{1}{wR(t)-\phi A}\right\}. \qquad (18)\end{aligned}$$

Solving for the point $\partial \mathbf{E}U\left(A,0\right)/\partial A = 0$ gives

$$A^* = \arg\max_A \mathbf{E}U\left(A,0\right) = \frac{pR\left(t\right)}{1-\phi}\left[R\left(p\right)R\left(\phi\right)-1\right]w. \qquad (19)$$

Rewriting A^* in (19) as $A^* = wR(t)\left[1-p-\phi\right]\phi^{-1}\left[1-\phi\right]^{-1}$, substituting this expression for A^* into (1), and differentiating with respect to E gives

$$\begin{aligned}\frac{\partial \mathbf{E}U\left(A^*,E\right)}{\partial E} &= \frac{[1-p]\phi}{[1-p]wR(t)+E\phi} - \frac{pf\left[1-\phi\right]}{pwR(t)-Ef\left[1-\phi\right]} \qquad (20)\\ &= \phi p\left\{\frac{R\left(p\right)}{[1-p]wR(t)+E\phi} - \frac{fR(\phi)}{pwR(t)-Ef\left[1-\phi\right]}\right\} \qquad (21)\end{aligned}$$

Evaluating at $\partial \mathbf{E}U\left(A^*,E\right)/\partial E = 0$ and solving for E gives

$$E^* = \arg\max_E \mathbf{E}U\left(A^*, E\right) = \frac{pR(t)}{1-\phi}\left[\frac{[1-p]\left[1-fR(\phi)\right]}{f}\right]w. \tag{22}$$

From (19) and (22) we see that

$$A^* > 0 \iff R\left(p\right)R\left(\phi\right) > 1; \qquad E^* > 0 \iff fR(\phi) < 1. \tag{23}$$

From (19) and (22) we compute

$$A^* + E^* = \frac{pR(t)}{1-\phi}\left[\frac{[1-p]\left[1-fR(\phi)\right] + f\left[R\left(p\right)R\left(\phi\right)-1\right]}{f}\right]w, \tag{24}$$

so

$$A^* + E^* < w \iff \frac{pR(t)}{1-\phi}\left[\frac{R(p) - f + fR(p)\left[R(p)R(\phi)-1\right]}{f}\right] < 1. \tag{25}$$

Proof of Proposition 5: Noting that

$$\left.\frac{\partial^2 E^*}{\partial p \partial f}\right|_{\mu_E = cons.} = -\frac{wR(t)\left\{2\left[1+f\right]R(p) - f\right\}}{f^3\left[1-\phi\right]\left[1+R(p)\right]} < -\frac{wR(t)\left[1+2f\right]}{f^2\left[1-\phi\right]\left[1+f\right]} < 0, \tag{26}$$

it follows that if $\partial E^*/\partial p|_{\mu_E=cons.} > 0$ when f approaches its maximum possible value, i.e., $f \to R\left(\phi\right)^{-1}$, then $\partial E^*/\partial p|_{\mu_E=cons.} > 0$ for all $f < R\left(\phi\right)^{-1}$. At $f \to R\left(\phi\right)^{-1}$ we indeed have

$$\left.\frac{\partial E^*}{\partial p}\right|_{\mu_E = cons.} = \frac{wR(p)R(t)R(\phi)\left[1+R(\phi)\right]}{\left[1-\phi\right]\left[1+R(p)\right]} > 0. \tag{27}$$

Chapter 3

Optimal Income Tax Enforcement in the Presence of Tax Avoidance

3.1 Abstract

We examine the optimal auditing problem of a tax authority when taxpayers can choose both to evade and avoid. For a convex penalty function the incentive-compatibility constraints may bind for the richest taxpayer and at a positive level of both evasion and avoidance. The audit function is non-increasing in reported income, and is higher for progressive tax functions than for regressive tax functions. Higher marginal tax rates increase the incentives for non-compliance, overturning the well-known Yitzhaki paradox.

Acknowledgements: We are grateful to Andrea Vindigni, Chris Hutcheon; and participants at the TARC Workshop on Tax Avoidance (London) and the 4th Annual TARC Workshop (Exeter) for helpful comments. This chapter was written while Gamannossi degl'Innocenti was a visitor at Brunel University London, whose hospitality he thanks. Gamannossi degl'Innocenti gratefully acknowledges financial support from the Ministero dell'Istruzione, dell'Università e della Ricerca (cycle XXVIII) and from the European Com-

3.2 Introduction

Individuals take a variety of actions to reduce their tax liabilities. In particular, one may distinguish between actions that are clearly in breach of the law (tax evasion); actions that are not explicitly ruled out under law, but which violate its spirit (tax avoidance); and actions that are legitimate (tax planning). The term tax avoidance is, however, sometimes used to refer to any action that changes a tax liability purely by affecting the form (but not the level) of compensation. For instance, tax avoidance opportunities arise in the context of income tax when taxpayers can shift part of their taxable income into profit or into another time period that is treated more favourably from a tax perspective-for a detailed discussion of these form-changing actions see, e.g., Stiglitz (1986) and Slemrod and Yitzhaki (2002). The type of tax avoidance we have in mind in this chapter is a narrower notion, in the sense that many form-changing actions are perfectly legal, and therefore fall into our notion of tax planning. Instead, we consider acts of form-changing that are so artificial in nature that the courts will deem them illegal if the tax authority mounts a legal challenge. These acts are often complex, and unlike evasion must be purchased from specialist providers known as promoters. A recent example of this type of avoidance scheme is a 2012 legal case in the UK between H.M. Revenue and Customs (HMRC) and a businessman named Howard Schofield. Schofield bought an avoidance scheme to help him reduce the amount of tax due on a 10m. capital gain on a share holding. The scheme used self-cancelling option agreements that would return the seller to his original position yet create an allowable

mission (Erasmus mobility grant 2015-1-IT02-KA103-013713/5).

loss. Although, viewed separately, the options created exempt gains and allowable losses, when viewed together as a composite transaction they did not. HMRC (2012) described the scheme as an artificial, circular, self-cancelling scheme designed with no purpose other than to avoid tax, and it was ultimately outlawed.

The first economic studies relating to tax compliance (e.g., Allingham and Sandmo 1972; Srinivasan 1973; Yitzhaki 1974; Christiansen 1980) utilised a general economic model of crime owing to Becker (1968). As this model lends itself much more readily to tax evasion (which is a crime) than tax avoidance (which is not outright illegal), these studies neglect the possibility of tax avoidance altogether. The economic literature that followed has largely retained this bias, even though in many countries it seems likely that loss of tax revenue due to avoidance activity is significant. For instance, according to Cobham (2005), developing countries lose $285 bn. per year due to tax evasion and tax avoidance. Estimates provided by the UK tax authority put the value of tax avoidance at £2.7 bn., compared to £4.4 bn. for tax evasion HMRC (2015). Lang et al. (1997) estimate that tax avoidance costs the German exchequer an amount equal to around 34% of income taxes paid.

One of the chief lines of enquiry of economists has been to study how a tax authority can collect a given amount of income tax revenue at minimum enforcement cost, when taxpayers can illegally under-report their true income. The instruments potentially available to the tax authority to achieve this objective are (i) a tax function, which associates a tax liability to each level of income; (ii) a penalty function, which associates a level of penalty to each level of evaded tax, and (iii) an audit function, which associates a probability of audit to each level of reported income.

Like much of the literature, we focus on the audit function by exogenously assuming the form of the penalty and/or tax functions. This is justified if (i) the entity that sets the audit function (the tax authority) does not have discretion over fiscal policy and (ii) the setting of penalties is highly constrained. In practice both these conditions usually hold: the design of the tax function is typically seen as a *policy* matter to be determined by the Treasury (whereas the collection of tax is seen as an *operational* matter), while the penalty function is fixed in legislation (making it costly to change) and is bounded in its severity by the requirement that it be proportional to the perceived seriousness of the crime. Sanchez and Sobel (1993) assume that taxpayers are risk neutral, that the penalty rate on undeclared tax is constant, and that the tax function is given. They give general conditions under which tax revenue is most efficiently collected as follows: taxpayers reporting an amount of income above a threshold amount are audited with probability zero, while taxpayers reporting an amount of income below the threshold are audited with a probability that is just sufficient that they will choose to report their income truthfully.[1] Given this audit probability function, all taxpayers with true income above the threshold amount declare exactly the threshold amount, and so pay the same amount of tax. Accordingly, the "effective" tax function (after taking into account the non-payment of tax due to under-reporting) becomes flat above the threshold declaration amount.

Another strand of literature assumes that a unified entity can simultaneously set the audit, penalty, and tax functions. In this setting Chander

[1]Earlier contributions that arrived at the conclusion of an audit threshold under less general assumptions include Reinganum and Wilde (1985), Scotchmer (1987) and Morton (1993) .

and Wilde (1998) show that, if taxpayers are risk neutral and fines are maximal, then the effective tax function is regressive and the audit function is non-increasing. Marhuenda and Ortuño-Ortn (1994) show that these results continue to hold for a range of other (seemingly more reasonable) penalty functions. Chander (2007) generalises these results to a particular class of risk averse preferences.[2] Few other general results exist, however: for instance, Mookherjee and Png (1989) show that the introduction of risk aversion can imply that the audit function is not always non-increasing in the amount of income declared.

In this chapter we investigate how accounting for the ability of individuals to avoid tax, as well as to evade tax, alters the conclusions of models in which only tax evasion is possible. In our model individuals can engage in tax evasion by under-reporting their income, but can also, at a cost, participate in a tax avoidance scheme that permits them to further lower reported income. Additional to the financial cost of avoidance, both forms of non-compliance are assumed, when detected, to impose psychic harm in the form of a social stigma cost. The nature of the avoidance scheme is not unambiguously prohibited by law, but is unacceptable to the tax authority. Accordingly, if the tax authority learns of the scheme, it will move to outlaw it ex-post. If a taxpayer is audited the tax authority observes whether they are using a tax avoidance scheme and also the extent of any tax evasion. The taxpayer is fined on the evaded tax, but the tax authority has no grounds to impose a fine on the avoided tax (it can only take measures to outlaw the scheme and then recover the tax owed on the avoided income). In this context we char-

[2] See, however, Hindriks (1999) for situations in which the regressivity of the tax function is reversed.

acterise the audit function first for a linear penalty function, and later for a general penalty function. The tax authority can condition its audit function only on the amount of income declared; it does not observe the amount of non-compliance or how it is split between evasion and avoidance. We therefore look for a taxpayer with income w^* and a level of tax avoidance A^* such that, if this taxpayer (weakly) prefers to report truthfully rather than hide an amount of income A^* then all other taxpayers will also wish to report truthfully.

We find that, if the penalty function is linear or strictly concave then, irrespective of the tax function, it holds that (i) if the wealthiest taxpayer is induced to report honestly, so will all other taxpayers; and (ii) at every income declaration, x, enforcement must be just sufficient that the wealthiest taxpayer does not wish to evade the amount of income $w - x$ (if evasion is more attractive than avoidance), or does not wish to avoid the amount of income $w - x$ (if avoidance is more attractive than evasion). That is, if the tax authority enforces to the point where pure evasion/avoidance becomes unattractive then mixtures of evasion and avoidance will also be unattractive. On the other hand, if the penalty function is convex (the marginal rate of penalty is increasing) then it is possible that the focus of enforcement is not the wealthiest taxpayer, but rather a taxpayer with intermediate wealth. The level of wealth of this critical taxpayer is an increasing function of income declared, implying that the focus of enforcement is on lower wealth individuals at lower levels of declared income, and on higher wealth individuals at higher levels of declared income. It also becomes possible that taxpayers prefer engaging simultaneously in evasion and avoidance over pure strategies. When this is so the optimal mix of avoidance and evasion moves in favour of avoidance as reported income decreases, as the competitiveness of the market for avoidance schemes increases, and as the social stigma associated with tax non-compliance falls.

In all cases we find the audit function to be a non-increasing function of declared income. When enforcement is predicated on the wealthiest taxpayer the audit function is strictly decreasing in declared income. The function is shifted upwards by an increase in wealth (of the wealthiest taxpayer), and shifted downwards by a steepening of penalties, an increase in the social stigma attached to tax non-compliance, and a lessening of competition in the market for avoidance schemes. When the focus of enforcement is not the wealthiest taxpayer, however, the audit function becomes independent of declared income and of the competitiveness of the market for avoidance. By analysing the audit function under example progressive and regressive tax functions we find that, as in Chander and Wilde (1998), lower enforcement is required to enforce a regressive tax than to enforce a progressive tax. Stronger risk aversion moves the audit function downwards, with larger downward movements for lower values of reported income.

We also find that an increase in marginal rates of tax stimulates incentives for non-compliance, such that the audit function must rise to maintain truthful reporting. This is the opposite finding of Yitzhaki (1974), in which the incentives to be non-compliant diminish as marginal tax rates increase. The difference in predictions is of interest as Yitzhaki's finding is counter-intuitive and at variance with most empirical evidence. Whereas taxpayers can only evade in Yitzhaki's model, in our model they can also avoid.We find that the incentives to avoid unambiguously increase following an increase in marginal tax rates, so even though the incentives for evasion may worsen, nonetheless the tax system becomes more costly to enforce, and compliance falls unless enforcement is stiffened.

The chapter adds to the small, but growing, economic literature on tax avoidance (in the broad sense). Slemrod and Kopczuk (2002), Piketty et al. (2014) and Uribe-Teran (2015) analyse theoretically the elasticity of taxable income in the presence of avoidance, while Alm (1988b) ,Alm and McCallin (1990) and Alm et al. (1990) examine the choice of an individual between evasion and avoidance. In the empirical literature, Slemrod (1995, 1996) finds pronounced tax avoidance effects in the response of high-earners to tax changes, while Feldstein (1999) finds that accounting for tax avoidance significantly increases estimates of the implied deadweight loss of income taxation. Fack and Landais (2010) show that the response of charitable deductions to tax rates is concentrated primarily along the avoidance margin (rather than the real contribution margin), while Gruber and Saez (2002) show that the elasticity of a broad measure of income is notably smaller than the equivalent elasticity for taxable income, suggesting that much of the response of taxable income comes through deductions, exemptions, and exclusions.

The plan of the chapter is as follows: section 3.3 outlines the model; section 3.4 performs the main analysis; and section 3.5 considers a range of extensions. Section 3.6 concludes. All proofs are in the Appendix.

3.3 Model

A taxpayer has an income (wealth) w; w differs among individuals on the support $[0, \overline{w}]$, where $\overline{w} > 0$. Each taxpayer faces a tax on income

w given by $t(w)$, satisfying $t(w) < w$ and $t' \geq 0$. Taxpayers behave as if they maximise expected utility, where utility is denoted by $U(z) = z$ (risk neutrality). A taxpayer's true income w is not observed by the tax authority, but the taxpayer must declare an amount $x \in [0, w]$. A taxpayer can choose to illegally evade an amount of income E and to avoid paying tax on a further amount of income A, where $x = w - E - A$.

Evasion is financially costless but avoidance technology is bought in a market in which "promoters" sell avoidance schemes to "users".[3] A common feature of this market is the "no saving, no fee" arrangement under which the price received by a promoter is linked to the amount by which they are able to reduce the user's tax bill. Although systematic information regarding the precise contractual terms upon which avoidance schemes are typically sold is scarce, we understand from a detailed investigation in the UK that, for the majority of mass-marketed schemes, the fee is related to the reduction in the annual theoretical tax liability of the user, not the ex-post realization of the tax saved (Committee of Public Accounts 2013). This implies, in particular, that the monetary risks associated with the possible subsequent detection and termination of a tax avoidance scheme are borne by the user [4]

Accordingly, we assume that the promoter's fee is a proportion $\phi \in (0, 1)$ of the tax saving accruing from the Scheme. In this way, ϕ may be interpreted as measuring the degree of competition in the market for tax avoidance schemes, with lower values of ϕ indicating the presence

[3]For analyses of the market for tax advice see, e.g., Reinganum and Wilde (1991) Damjanovic and Ulph (2010).

[4]It is apparent that such arrangements give promoters incentives to mis-represent the level of risk involved in particular schemes. Consistent with this point, (Committee of Public Accounts, 2013, p.11) indeed finds evidence of such mis-selling.

of stronger competitive forces. When a taxpayer is simultaneously evading and avoiding, the tax saving accruing to the avoidance scheme is not always unambiguous, however. To see this, note that the total tax underpayment of a taxpayer is given by $t(w) - t(x)$. This can be decomposed in two ways: one decomposition is to assign $t(w) - t(w - E)$ to be the evaded tax, and $t(x + A) - t(x)$ to be the avoided tax, but an alternative taxonomy is to assign $t(x + E) - t(x)$ to be the evaded tax and $t(w) - t(w - A)$ to be the avoided tax. These alternative approaches are equivalent if the tax function is assumed to be linear, but are distinct otherwise. As our results are not especially sensitive to which of these conventions is adopted, however, we adopt the first of these decompositions in our baseline specification. Hence, we may write the total fee paid by the taxpayer to the promoter as $\phi [t(x + A) - t(x)]$.

We adopt a principal-agent approach in which the principal can commit to an audit and penalty function which taxpayers then take as given. Though important, as in many other contexts, we do not address the issue of how the principal can make these commitments.[5] A taxpayer reporting income x is audited with probability $p(x)$. If audited, E and A are observed. A taxpayer must then make a payment $f(t(w) - t(w - E))$ on account of the amount of evaded tax, where $f(0) = 0$ and $f' > 1$ (which, together, imply $f(z) > z$ for $z > 0$). The taxpayer cannot be fined on the avoided tax, however. The tax authority mounts a (successful) legal challenge to the avoidance scheme, giving the tax authority the right to reclaim the tax owed. Thus, instead of paying $t(x)$, the taxpayer must pay $t(x + A)$.

[5]Reinganum and Wilde (1986) and Erard and Feinstein (1994) study the case of the principal not being able to make commitments.

The experiments of Baldry (1986) provide compelling evidence that the non-compliance decision is not just a simple gamble. This can be rationalized by introducing an additional cost into the decision. This cost can be financial (Chetty 2009; Lee 2001) or psychic. We adopt a psychic cost interpretation, where the psychic cost is identified as the social stigma associated with being caught performing activities that either abuse the spirit of the law, or outright violate it. Other models to allow for costs due to social stigma include Al-Nowaihi and Pyle (2000), Benjamini and Maital (1985), Dell'Anno (2009), Dhami and Al-Nowaihi (2007), Gordon (1989), and Kim (2003). Social stigma is incurred when $A + E \ (= w - x) > 0$ and the taxpayer is audited. Specifically, we write

$$S\left(w - x\right) = \begin{cases} 0 & \text{if } x = w; \\ s > 0 & \text{otherwise.} \end{cases}$$

One might think that the stigma cost, as well as having a fixed component, might also have a component that increases in the total amount of non-compliance ($A + E$). We shall allow for this possibility in Section 3.5 as an extension to the baseline model.[6] It might also be argued that the social stigma associated with avoidance and evasion differ. For instance, Kirchler et al. (2003) find socially positive attitudes towards tax avoidance (but socially negative attitudes towards tax evasion) among students, fiscal officers and small business owners in Austria. Recent poll evidence for the UK, however, suggests that evasion and avoidance are viewed similarly (Stone 2015). Given the mixed evidence, and that public attitudes may well vary over time, assuming that social stigma is

[6] A further line of literature (see, e.g., Hashimzade et al. 2014, 2015, 2016; Myles and Naylor 1996 relates social stigma to the prevalence of non-compliance among taxpayers. We do not explore this route here, but offer it as a possible avenue for research.

associated equally with avoidance and evasion seems reasonable.

A taxpayer's expected utility is therefore given by

$$EU = [1 - p(x)]\, U^n + p(x)\, U^a, \tag{1}$$

where U^n is a taxpayer's utility in the state in which they are not audited and U^a is a taxpayer's utility in the state in which they are audited. We then write $U^n \equiv U(w^n)$ and $U^a \equiv U(w^a) - S(w - x)$, where $\{w^a, w^n\}$ are, respectively, a taxpayer's wealth in the audit and non-audit states. Note that, owing to the equality $x = w - E - A$, we can write w^a and w^n as either functions of $\{x, A, w\}$ or of $\{E, A, w\}$. As each formulation yields separate insights we define both here. In the former case we have

$$\begin{aligned} w^n(x, A, w) &= w - t(x) - \phi[t(x + A) - t(x)] \\ w^a(x, A, w) &= w - t(x + A) - f(t(w) - t(x + A)) - \phi[t(x + A) - t(x)] \end{aligned}$$

and in the latter we have

$$\begin{aligned} w^n(A, E, w) &= w - t(w - A - E) - \phi[t(w - E) - t(w - A - E)] \\ w^a(A, E, w) &= w - t(w - E) - f(t(w) - t(w - E)) - \\ &\qquad \phi[t(w - E) - t(w - A - E)] \end{aligned}$$

We adopt the standard assumption of *limited liability*, whereby the tax and fine payments of a taxpayer cannot exceed their wealth w. Accordingly, to ensure that the limited liability condition always holds, we assume $w^a(x, A, w) - s > 0$, a necessary condition for which is that $w - s \geq f(t(w))$.

A *mechanism* for the tax authority consists of a set of possible income reports $M \in [0, w]$, a tax function $t(\cdot)$, an audit function $p(\cdot)$, and a penalty function $f(\cdot)$. In this chapter we focus only on *incentive compatible* mechanisms, i.e., mechanisms that induce all taxpayers to report truthfully. The standard justification for this approach is the *revelation principle*: when this holds then, for any feasible mechanism, one can find an equivalent mechanism that induce taxpayers to report truthfully (see, e.g., Myerson 1979, 1982, 1989). Chander and Wilde (1998) show that the revelation principle applies when the tax authority has unfettered ability to choose the tax and audit functions, while the penalty function is only constrained to be bounded above. Unfortunately, penalty functions of this type deviate significantly from those observed in practice as the penalty for under-reporting by any amount, no matter how small, is extreme. As noted by Cremer and Gahvari (1995), however, adopting more appealing but exogenously given penalties implies that one can no longer rely on the revelation principle. Whereas most of the literature has implicitly opted for tractability over realism, here we follow the lead of Marhuenda and Ortuño-Ortn (1994) in considering a setting in which the revelation principle does not hold. Implicitly, therefore, we restrict attention to the set of mechanisms that are payoff equivalent to the set of incentive compatible mechanisms we consider here. Our focus shall be primarily on the shape of the audit function for a given penalty and tax function. Accordingly, we do not allow the tax authority to choose the latter two functions.

The utility when reporting truthfully (honestly) is $U^h \equiv U\left(w^h\right)$, where $w^h = w - t(w)$. In order that the mechanism be incentive compatible, a taxpayer must never receive a utility higher than $U\left(w^h\right)$ when reporting $x < w$. This implies that

$$p(x) \geq \frac{U^n - U^h}{U^n - U^a} \text{ for all } A \in [0, w-x],\, x \in [0, w] \text{ and for all } w. \quad (2)$$

Performing an audit costs the tax authority an amount $c > 0$. Given this, a revenue maximising scheme will always minimise $p(x)$ subject to the condition in (2) holding. Define the function $p(x; A, w)$ as the smallest probability of audit that induces an (A, w)-taxpayer to report truthfully. Then

$$p(x; A, w) = \begin{cases} \frac{U^n - U^h}{U^n - U^a} = \frac{t(w) - t(x) - \phi[t(A+x) - t(x)]}{f(t(w) - t(A+x)) + t(A+x) - t(x) + s} & x < w; \\ 0 & x = w. \end{cases} \quad (3)$$

The restriction $p(x; A, w) \leq 1$ holds necessarily as $U^h \geq U^a$. When $x = w$ the definition of $p(x; A, w)$ becomes arbitrary, for the condition in (2) must hold for any $p(x)$. In setting $p(w; A, w) = 0$ we follow (Chander, 2007, p.325). In what follows we restrict attention to the case $x < w$ unless it is explicitly stated otherwise. Note in (3) that the tax function always appears in the form $t(z_1) - t(z_2)$, with the implication that the audit function is independent of the *level* of the tax function (any vertical shift of $t(\cdot)$ must cancel). Accordingly, it is without loss of generality that we set $t(0) = 0$.

The tax authority cannot, however, utilise $p(x; A, w)$ as it observes x, but not A or w. Instead, the tax authority must choose $p(x)$ such that, for each x, reporting is truthful for all feasible A and w. Accordingly, we then define $p(x)$ as

$$p(x) = \max_{A, w} p(x; A, w)$$

3.4 Analysis

We begin by considering the special case in which taxpayers are risk neutral ($U'' = 0$). We begin without restricting the form of the tax function, but restrict the penalty function to be linear: $f(z) = [1+h]\,z$, $h > 0$. In this way we obtain a very simple version of the model that provides ready intuitions. For each value of x, we wish to maximise $p(x; A, w)$ in (3) with respect to A and w (allowing the suppressed variable E to vary). First, maximising with respect to A, the first order condition for a maximum is

$$\frac{\partial p(x; A, w)}{\partial A} = -\frac{\{\phi s + \{\phi - h[1-\phi]\}[t(w) - t(x)]\}\, t'(A+x)}{\{[1+h][t(w) - t(A+x)] + t(A+x) - t(x) + s\}^2} \quad (4)$$

Define $A^* = \arg\max_A p(x; A, w)$ and $w^* = \arg\max_w p(x; A, w)$. Then (4) implies that $A^* = 0$ when

$$\phi > \hat{\phi} = \frac{h[t(w) - t(x)]}{s + [1+h][t(w) - t(x)]}$$

and $A^* = w - x$ when $\phi < \hat{\phi}$. When $\phi = \hat{\phi}$ all feasible values of A weakly maximise $p(x; A, w)$. Taking the case $\phi > \hat{\phi}$ first, to find $p(x)$ we now maximise $p(x; 0, w)$ with respect to w. The first derivative with respect to w is

$$\frac{\partial p(x; 0, w)}{\partial w} = \frac{s t'(w)}{\{[1+h][t(w) - t(x)] + s\}^2} > 0 \quad (5)$$

so $w^* = \overline{w}$. In the case $\phi < \hat{\phi}$ the relevant first derivative with respect to w is

$$\frac{\partial p(x; w-x, w)}{\partial w} = \frac{s[1-\phi]\, t'(w)}{[t(w) - t(x) + s]^2} > 0 \quad (6)$$

	$A^*=0$	$A^*=w^*-x$	$A^*\in(0,w^*-x)$		$w^*\in(x+A^*,\overline{w})$	
	$p(x)$	$p(x)$	A^*	$p(x)$	w^*	$p(x)$
x	$-$	$-$	$-$	$-$	$+$	0
$\overline{w}$	$+$	$+$	$+$	$+$	0	0
ϕ	0	$-$	$-$	$-$	0	0
s	$-$	$-$	$-$	$-$	$+$	$-$
pivot $f(\cdot)$	$-$	0	$+$	$-$	$-$	$-$
pivot $t(\cdot)$	$+$	$+$	$+$	$+$	$-$	0

Table 1: Comparative statics

so again $w^*=\overline{w}$.

Proposition 6 *If the penalty function is linear then*

$$p\left(x\right)=\begin{cases}\frac{[1-\phi][t(\overline{w})-t(x)]}{t(\overline{w})-t(x)+s} & \phi<\hat{\phi};\\ \frac{t(\overline{w})-t(x)}{f(t(\overline{w})-t(x))+s} & \phi>\hat{\phi}.\end{cases} \tag{7}$$

Summarising this analysis, when the market for avoidance schemes is sufficiently competitive ($\phi<\hat{\phi}$) it is sufficient to incentivise truthful reporting by all taxpayers that the wealthiest taxpayer does not wish to avoid all of their income. This holds irrespective of the shape of the tax function. If, however, $\phi>\hat{\phi}$ then evasion is more attractive to taxpayer's than is avoidance. In this case it is sufficient to incentivise truthful reporting that the wealthiest taxpayer does not wish to evade all of their income.

The form of $p\left(x\right)$ in (7) applies more generally whenever A^* takes corner values and $w^*=\overline{w}$ (not only when the penalty function is linear). It transpires that a corner solution necessarily arises when $f''\leq 0$, and may also arise when $f''>0$ under further conditions. We now analyse the comparative statics properties of $p\left(x\right)$ in (7).

Proposition 7 *In an equilibrium in which $A^*\in\{0,w-x\}$ and $w^*=\overline{w}$ then the comparative statics of $p\left(x\right)$ are given as in columns 1 and 2 of Table 1.*

Proposition 7 is most readily understood with respect to the expected marginal returns to evasion and avoidance. The expected return to the gamble of reporting $x < w$ (rather than w) is given, for a fixed p, by

$$R(A, E) = p\left[w^c(A, E, w) - s\right] + (1 - p)w^n(A, E, w) - w^h(w) \tag{8}$$

In the formulation in (8) we retain A and E, by suppressing x. This allows us to consider, e.g., the effect of moving A holding E constant (with x adjusting to maintain the equality $x = w - E - A$). As taxpayers are risk neutral it must hold that $R(A^*, E^*) = 0$, for if $R(A^*, E^*) > 0$ incentive compatibility is violated, and if $R(A^*, E^*) < 0$ the tax authority could achieve truthtelling at lower cost. From (8) the expected marginal benefit to, respectively, E and A (for a fixed p) are therefore given by

$$\frac{\partial R}{\partial A} = (1 - p - \phi)t'(w - A - E) \tag{9}$$

$$\frac{\partial R}{\partial E} = \frac{\partial R}{\partial A} - \{p[f' - 1] - \phi\}\, t'(w - E) \tag{10}$$

The corner solution $A^* = 0$ arises when $\partial R/\partial E > \partial R/\partial A$ for all A and the corner solution $A^* = w - x$ when $\partial R/\partial A > \partial R/\partial E$ for all A. As the $p(x)$ in Proposition 7 is predicated on requiring the wealthiest taxpayer to report truthfully, it is responsive to changes in $\overline{w}$. In particular, when $A^* = 0$, if the wealthiest taxpayer chooses to evade in full an incremental increase in their income, the effect on the expected return to evasion is given by

$$\left.\frac{\partial R}{\partial \overline{w}}\right|_{x=\text{const.}} = \left[1 - pf'\left(t(\overline{w}) - t(\overline{w} - E)\right)\right] t'(\overline{w})$$

Note by inspection of (7) that at the corner solution $A^* = 0$ it holds that $p < \left[f'\left(t\left(\overline{w}\right) - t\left(x\right)\right)\right]^{-1}$, so $1 - pf'\left(t\left(\overline{w}\right) - t\left(\overline{w} - E\right)\right) > 0$. It follows that $\partial R / \partial \overline{w}|_{x=\text{const.}} > 0$, so the probability of audit must necessarily rise to maintain a zero expected return to non-compliance. If instead $A^* = w - x$ then, if the wealthiest taxpayer chooses to avoid in full an incremental increase in their income, the effect on the expected return to avoidance is given by

$$\left.\frac{\partial R}{\partial \overline{w}}\right|_{x=\text{const.}} = \left[1 - p - \phi\right] t'\left(\overline{w}\right)$$

By inspection of (7), at the corner solution $A^* = w - x$ it holds that $p < 1 - \phi$, so necessarily $\partial R / \partial \overline{w}|_{x=\text{const.}} > 0$. Again, the probability of audit must rise to preserve a zero expected return. Hence, whichever corner solution for A applies, the audit function is increasing in the wealth of the wealthiest taxpayer. As it is gainful to the wealthiest taxpayer to increase evasion (when $A^* = 0$) and avoidance (when $A^* = w - x$) it follows that to discourage the taxpayer from reporting low values of x requires more enforcement activity than does discouraging the reporting of higher values, hence the audit function is decreasing in reported income.

When the avoidance market is sufficiently competitive that avoidance is a superior instrument in reducing a taxpayer's liability than is evasion (i.e., $\partial R / \partial A > \partial R / \partial E$) a further increase in the competitiveness of the market for avoidance schemes (a fall in ϕ) induces the wealthiest taxpayer to wish to avoid more, and forces $p(x)$ to increase to maintain truth-telling. When, however, the avoidance market is sufficiently uncompetitive that in any case avoidance is unappealing (relative to evasion) as a means of reducing one's tax liability, then the audit function becomes independent of ϕ. Similarly, a multiplicative shift in the penalty function (which increases the marginal rate of penalty by a fixed proportion) only affects $p(x)$ when the wealthiest taxpayer wishes to evade rather than avoid ($A^* = 0$). In this case evasion becomes more costly at the margin, thereby relaxing the truth-telling constraint. We also see that an increase in social stigma results in a fall in the attractiveness of both evasion and avoidance, allowing $p(x)$ to fall while maintaining honest reporting.

A proportional increase in marginal tax rates (a multiplicative shift of the tax function such that $t(\overline{w}) - t(x)$ increases for every x) increases both the expected benefits and costs of evasion and avoidance, making its effect difficult to anticipate with intuition alone. In the absence of avoidance it is well-known that the standard model of tax compliance of Yitzhaki (1974) predicts that an increase in the marginal tax rate decreases the incentive to evade, which implies (in a model without avoidance) that the tax authority would therefore be able to lower the audit function while still achieving truthful reporting. In columns 1 and 2 of Table 1 we observe the opposite result: as marginal tax rates increase the audit function increases. To understand this result, first consider the corner solution $A^* = 0$. Here what is crucial is how the expected return to evasion responds to a multiplicative shift of the tax function. As $t(0) = 0$ a multiplicative shift can equally be thought of as an anti-clockwise pivot of $t(\cdot)$ around the origin (intercept). Hence we may write $t(\cdot)$ as $\varepsilon t(\cdot)$, and then consider $\lim_{\varepsilon \to 1} \partial R / \partial \varepsilon |_{A=0}$:

$$\lim_{\varepsilon \to 1} \left. \frac{\partial R}{\partial \varepsilon} \right|_{A=0} = \left[t'\left(w\right) - t'\left(w - E\right) \right] \left[1 - pf'\left(t\left(w\right) - t\left(x\right)\right) \right] > 0$$

Hence, when $A^* = 0$ evasion is made more attractive by stiffening marginal tax rates. When $A^* = w - x$ it is instead crucial how the expected return to avoidance responds to a multiplicative shift of the tax function. We have

$$\lim_{\varepsilon \to 1} \left. \frac{\partial R}{\partial \varepsilon} \right|_{A=w-x} = \left[1 - p - \phi \right] \left[t\left(w\right) - t\left(w - A\right) \right] > 0, \tag{11}$$

which implies that the audit function must shift upwards to restore the expected return to zero. Noting from (9) that $1-p-\phi > 0$ is the condition for avoidance to be gainful in expectation, (11) implies that, when avoidance is gainful in expectation, a multiplicative shift of the tax function will increase the expected return to avoidance.

Having established that a linear penalty function always leads to a corner A^* we now examine the case in which the penalty function is kept general. In particular, we are interested in understanding the conditions under which $A^* \in (0, w - x)$. An alternative approach to differentiating $p(x; A, w)$ directly (as we did above) is to exploit the observation that $R\left(A^*, E^*\right) = 0$. The implicit function theorem (IFT) then implies that (13) and (14) can also be rewritten more generally as

$$\frac{\partial p(x; A, w)}{\partial z} = \frac{\left[\frac{\partial w^a}{\partial z} - \frac{\partial w^n}{\partial z} \right] p(x; A, w) + \frac{\partial w^n}{\partial z} - \frac{\partial w^h}{\partial z}}{w^n - w^a + s}; \qquad z \in \left\{ A, w \right\}, \tag{12}$$

giving

$$\frac{\partial p(x;A,w)}{\partial A} = \frac{\{p(x;A,w)\,[f'-1]-\phi\}\,t'\,(x+A)}{w^n-w^a+s} \tag{13}$$

$$\frac{\partial p(x;A,w)}{\partial w} = \frac{[1-p(x;A,w)f']\,t'\,(w)}{w^n-w^a+s} \tag{14}$$

Using (13), at a stationary point for A we have

$$p(x;A^*,w) = \frac{\phi}{f'-1} \tag{15}$$

and, from (14), at a stationary point for w we have

$$p(x;A,w^*) = \frac{1}{f'} \tag{16}$$

To verify when these define a maximum we use (13) and (14) to compute the second derivatives at a stationary point as

$$\left.\frac{\partial^2 p(x;A,w)}{[\partial A]^2}\right|_{\frac{\partial p(x;A,w)}{\partial A}=0} = -\frac{p(x;A,w)\,[t'\,(x+A)]^2\,f''}{w^n-w^a+s} \tag{17}$$

$$\left.\frac{\partial^2 p(x;A,w)}{[\partial w]^2}\right|_{\frac{\partial p(x;A,w)}{\partial w}=0} = -\frac{p(x;A,w)\,[t'\,(w)]^2\,f''}{w^n-w^a+s} \tag{18}$$

Inspecting equations (17) and (18) we see that their sign is the sign of f'', so for an interior maximum with respect to one or both of A and w it must hold that $f''>0$. We now investigate the case in which $A^*\in(0,w-x)$:

Lemma 1 *If* $A^*\in(0,w-x)$ *then* $p\,(x)\,f'<1<[1-\phi]\,f'$ *and* $p\,(x)<1-\phi$

In respect of the expected marginal returns to evasion and avoidance, Lemma 1 implies, first, that at an interior value of A^* the expected marginal return to evasion must equal the expected marginal return to avoidance:

$\partial R(A,E)/\partial A = \partial R(A,E)/\partial E$. Second, it implies that both expected marginal returns must be positive.

Using (13) and (14), the effect of w on $p(x;A,w)$ when $\partial p(x;A,w)/\partial A = 0$ is given by

$$\left.\frac{\partial p(x;A,w)}{\partial w}\right|_{\frac{\partial p(x;A,w)}{\partial A}=0} = \frac{[1-\phi-p(x;A,w)]\,t'(w)}{w^n - w^a + s} > 0$$

where the inequality follows from Lemma 1. This implies that when A^* is interior, w^* is maximal. Substituting $w = \overline{w}$ in (15) we therefore obtain

$$p(x) = \frac{\phi}{f'(t(\overline{w}) - t(x+A^*)) - 1} \tag{19}$$

From (13) and Lemma 1 we have

$$1 - p(x) - \phi = \frac{s + f(t(\overline{w}) - t(x+A^*)) - [t(\overline{w}) - t(x+A^*)]f'(t(\overline{w}) - t(x+A^*))}{s + f(t(\overline{w}) - t(x+A^*)) + [t(x+A^*) - t(x)]f'(t(\overline{w}) - t(x+A^*))} > 0$$

Hence, it must hold that $s > \varepsilon_f(t(\overline{w}) - t(x+A^*)) - 1$, where $\varepsilon_f(z) = zf'(z)/f(z)$ is the elasticity of the penalty function with respect to evaded tax, so interior values of A^* arise for sufficiently high social stigma costs.

These findings are illustrated in Figures 3, 4, and5.We depict $p(x)$ in Figure 3, the associated $\{A^*, E^*, w^*\}$ in Figure 4, and the expected marginal returns (denoted R_A and R_E for brevity) drawn at $p = p(x)$ and $E = E^*$ in Figure 5. [7]. For $x \in [0, \hat{x})$ A^* is interior – so $p(x)$ is as in (19). For $x \geq \hat{x}$ $A^* = 0$ – so $p(x)$ is as in Proposition 6.

We see in Figure 3 that $p(x)$ is decreasing and concave in x. Consistent with Lemma 1 we see that the audit function lies below $1/f'$, which

[7] Figures 3, 4, and5 are drawn for a linear tax function, $t(z) = 0.3z$, a quadratic penalty function of the form $f(z) = [1.1 + z/2]\,z$, $\phi = 0.2$, $s = 3$, and $\overline{w} = 10$

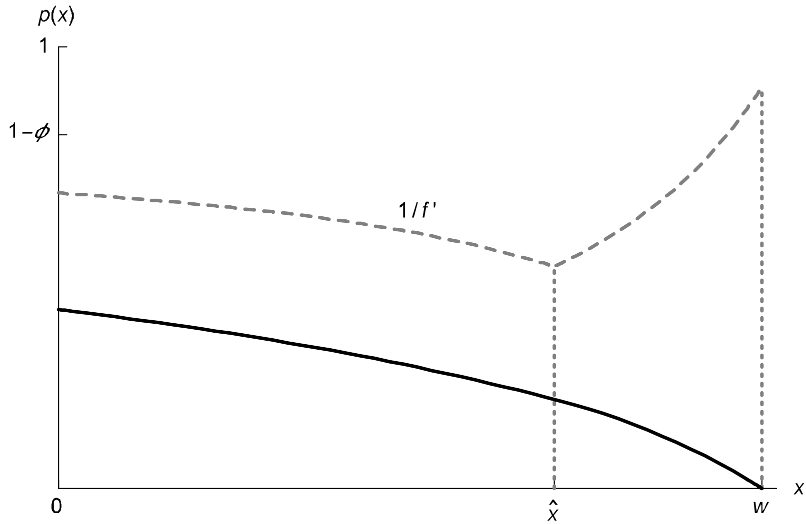


Figure 3: Audit function for $A^* \in (0, w^* - x]$.

is itself bounded above by $1 - \phi$. In Figure 4, A^* is initially decreasing and concave in x, and E^* is initially increasing and convex in x. In Figure 5 the expected marginal return to avoidance is seen to be constant in x. This is due to the choice of a linear fine rate; more generally, it is seen from (9) that tax avoidance displays increasing/constant/diminishing marginal returns as the tax function is regressive ($t'' < 0$)/linear ($t'' = 0$)/progressive ($t'' > 0$). To understand the shape of the expected marginal return to evasion, observe that the variation of the expected marginal return to evasion at different levels of evasion is given at the optimum by

$$\left.\frac{\partial^2 R}{\partial E^2}\right|_{\frac{\partial R(A,E)}{\partial A} = \frac{\partial R(A,E)}{\partial E}} = \frac{\partial^2 R}{\partial A^2} - p\left(x\right)\left[t'\left(w - E^*\right)\right]^2 f''$$

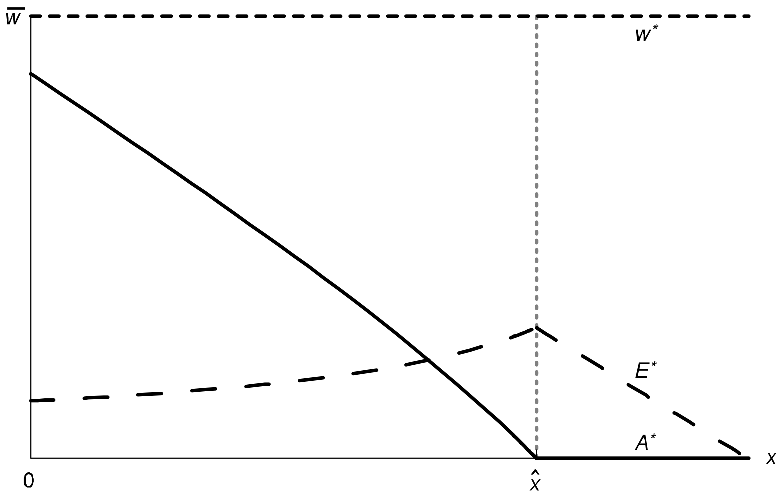


Figure 4: $\{A^*, E^*, w^*\}$ for $A^* \in (0, w^* - x]$.

As $f'' > 0$ at an interior A^*, it must hold that $\partial^2 R/\partial E^2 < \partial^2 R/\partial A^2$, as seen in the Figure.

We now consider the case in which $w^* \in (x + A, \overline{w})$. Proceeding in a manner similar to Lemma 1 we obtain that $p(x) = [f']^{-1} > 1 - \phi$. Referring to the expected marginal returns in (9) and (10), this inequality implies that w^* arises when $\partial R(A, E)/\partial E = 0 > \partial R(A, E)/\partial A$. At this point, the taxpayer does not wish to increase either evasion or avoidance at the margin.

Using (13) and (14), the effect of A on $p(x; A, w)$ when $\partial p(x; A, w)/\partial w = 0$ is given by

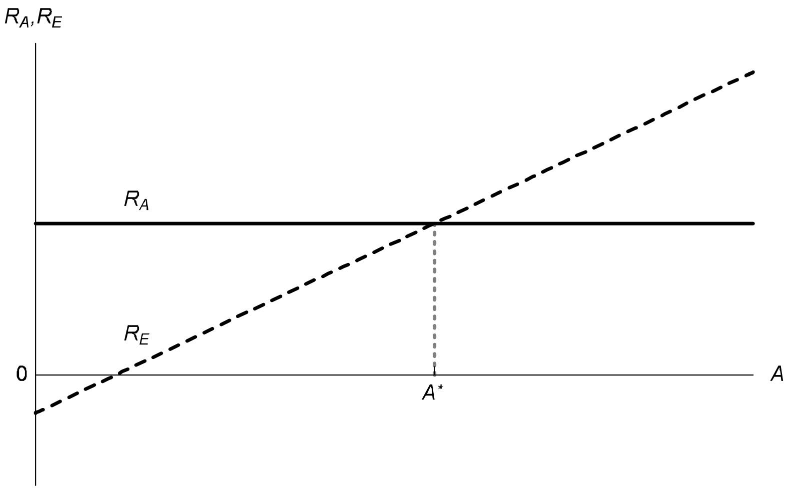


Figure 5: Expected marginal return to avoidance and evasion for $A^* \in (0, w^* - x)$.

$$\left.\frac{\partial p(x; A, w)}{\partial A}\right|_{\frac{\partial p(x;A,w)}{\partial w}=0} = [1 - p(x; A, w) - \phi]\, t'(x + A) < 0$$

This implies that when w^* is interior, A^* is minimal. Substituting $A = 0$ in (16) we therefore obtain

$$p(x) = \frac{1}{f'(t(w^*) - t(x))} \tag{20}$$

From (14) we have

$$1 - p(x) - \phi = 1 - \frac{[t(w^*) - t(x)]\, f'(t(w^*) - t(x))}{s + f(t(w^*) - t(x))} < 0 \tag{21}$$

As (21) is negative, it must be that $s < \varepsilon_f \left(t\left(w^*\right) - t\left(x\right)\right) - 1$. Hence, w^* is interior when a sufficiently low level of social stigma prevails, whereas A^* is interior when a sufficiently high level of social stigma prevails.

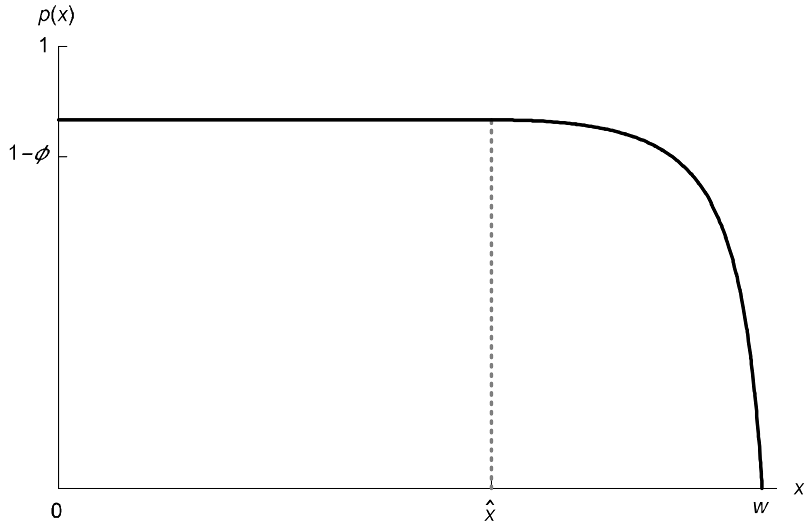


Figure 6: Audit function for $w \in (x + A^*, \overline{w}]$.

Our findings for the case when w^* is interior are illustrated in Figures 6, 7, and 8. Such figures are analogous to Figures 3, 4, and5, but to satisfy the condition in (21), we now decrease the stigma cost to $s = 0.1$. For $x \in [0, \hat{x})$ A^* is interior – so $p\left(x\right)$ is as in (20). For $x \geq \hat{x}$ $A^* = 0$ – so $p\left(x\right)$ is as in Proposition 6. In Figure 6 we see that $p\left(x\right)$ is initially independent of x, but falls rapidly in a concave manner after w^* reaches the upper bound $w^* = \overline{w}$. In this example $\partial w^* / \partial x = 1$ in Figure 7 but we shall show that, more generally, $\partial w^* / \partial x = t'\left(x\right) / t'\left(w\right)$. In Figure 8

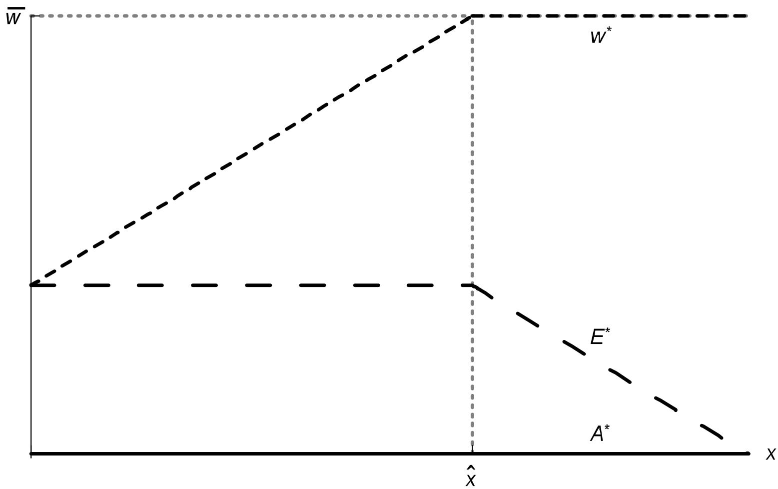


Figure 7: $\{A^*, E^*, w^*\}$ for $w^* \in (x + A^*, \overline{w}]$.

we see that the expected return to avoidance is negative for all w. The variation of the expected marginal return to evasion in w is given at the optimum by

$$\left.\frac{\partial^2 R}{\partial E \partial w}\right|_{\frac{\partial R(A,E)}{\partial E}=0} = -p\,(x)\,t'(w^*)t'(w - A - E)f''$$

As $f'' > 0$ at an interior w^*, it must hold that $\partial^2 R / \partial E \partial w < 0$, as seen in the Figure.

We now formally investigate the comparative statics of the two cases analysed above:

Proposition 8 *In an equilibrium in which either A^* or w^* takes an interior value the comparative statics of $\{A^*, p(x), w^*\}$ are given as in columns 3 and 4 of Table 1.*

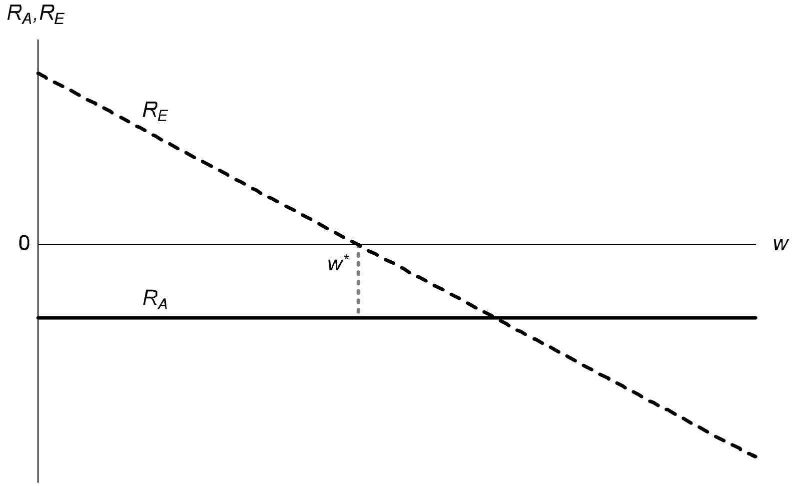


Figure 8: Expected marginal return to avoidance and evasion for $w^* \in (x + A^*, \overline{w})$.

When A^* takes an interior value the results in Table 1 (column 3) for the comparative statics of $p(x)$ are consistent with those obtained in Proposition 7: the audit function is a decreasing function of declared income, shifts downwards with increases in ϕ and s, and shifts upwards in $\overline{w}$. Moreover, $\partial A^* / \partial x$ can be written as

$$\frac{\partial A^*}{\partial x} = -1 - \frac{\{[1-\phi] f' - 1\} t'(x)}{[w^n - w^h] t'(x+A) f''} < -1,$$

with the implication that E^* is an increasing function of x (and A^*/E^* is a decreasing function of x). Whether A^*/E^* is an increasing or decreasing function of wealth depends on the shape of the tax function. If the tax function is progressive or linear then it can be shown that $\partial A^* / \partial \overline{w} > 1$, so

E^* must fall, but both A^* and E^* may rise if the tax function is regressive.

When w^* takes an interior value, however, the audit function becomes independent of declared income (and this holds for any tax function). The audit function also becomes independent of $\overline{w}$ (as it is not predicated on the wealthiest taxpayer) and of ϕ (as avoidance is dominated by evasion as a means of reducing tax liability). In both types of interior optimum a steepening of the penalty function shifts the audit function downwards.

We now return to the question of the effects of a proportional increase in marginal tax rates (a steepening of the tax function – again by means of an anti-clockwise pivot about the intercept). Matching our finding in Proposition 7 for the case of a corner solution, the findings in Table 1 predict the opposite of the Yitzhaki (1974) finding: as marginal tax rates increase the tax authority must raise the audit function to maintain truthful reporting. This finding is of note as Yitzhaki's result is not only paradoxical intuitively, but much empirical and experimental evidence finds a negative relationship between compliance and the tax rate (see, e.g., Bernasconi et al. 2014, and the references therein).[8] In interpreting this result it is of importance to note that the Yitzhaki (1974) model can be augmented with a constant utility cost due to social stigma – as in our model – without affecting the direction of the relationship between marginal tax rates and non-compliance.[9] This difference between models is not, therefore, a part of the explanation of our differing findings.

[8]See also Piolatto and Rablen (2016) for a detailed analysis of Yitzhaki's finding, and when it is and is not overturned.

[9]If, however, social stigma is viewed as a monetary, rather than utility cost, then a negative relationship between compliance and the marginal tax rate can emerge in the Yitzhaki framework when the stigma cost is sufficiently high (see, e.g., Al-Nowaihi and Pyle 2000).

Rather, the reversal of Yitzhaki's finding relies on the idea that, even in cases where evasion becomes less attractive following an increase in marginal tax rates, tax avoidance will become more attractive for sure. Thus the overall incentives for non-compliance grow, even if the incentives for evasion weaken.

We illustrate this point graphically in Figure 9, which shows the effect on the expected marginal returns to evasion (R_E) and avoidance (R_A) of a multiplicative shift of a (linear) tax function. Specifically, we increase the marginal tax rate from $t^- = 0.2$ to $t^+ = 0.7$ in the model specification used in Figures 3, 4, and5. The increase in marginal tax rates is seen to increase the expected marginal return to avoidance, so that the overall expected marginal return to non-compliance at the optimum is increased (making $p(x)$ higher). In this case the expected marginal return to evasion does not uniformly increase or decrease, but rather evasion becomes subject to stronger diminishing marginal returns (recall that evasion and avoidance are inversely related for a fixed x, so the amount of evasion increases from right to left in Figure 9).

3.5 Extensions

In this section we consider a range of realistic extensions to the model of the previous section. As, however, these extensions reduce (often substantially) the tractability of the model, we proceed here with solved examples, rather than general analytic solutions. As a key feature of our analysis is the incorporation of tax avoidance, we herein focus on the case in which the incentive compatibility constraints bind for an interior level of avoidance.

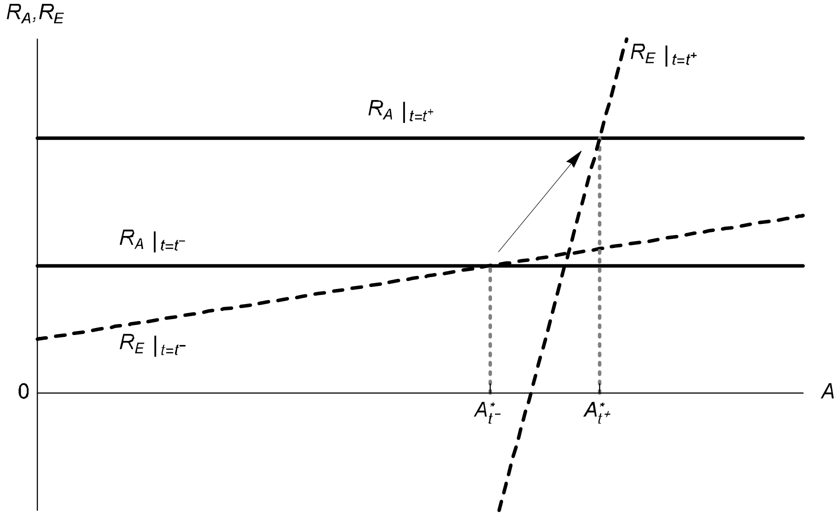


Figure 9: Effect of a multiplicative shift in the tax function on the expected marginal return to avoidance and evasion.

3.5.1 Optimal Auditing

We now revisit the finding of Chander and Wilde (1998) that regressive tax functions are more efficient than progressive tax functions (in the sense that they cost less to enforce). In Figure 10 we show $p(x)$ for the linear ($t'' = 0$), regressive ($t'' < 0$), and progressive ($t'' > 0$) cases.[10] As in previous Figures, A^* is interior for $x < \hat{x}$ and $A^* = 0$ for $x \geq \hat{x}$. We see that the audit function in the progressive case is everywhere above the audit function in the regressive case. Hence, the model retains Chander and Wilde's finding regarding the desirability of regressive taxation from an enforcement cost perspective. Our finding is not importantly altered if we instead employ the alternative formulation of the model whereby

[10]The specific functions depicted are $t(x) = 0.3x$ (linear case); $t\,(x) = 0.3x - 0.01x^2$ (regressive case); and $t\,(x) = 0.06x^2$ (progressive case).

$t\,(x+E) - t\,(x)$ is considered the evaded tax and $t\,(w) - t\,(w-A)$ is considered to be the avoided tax.

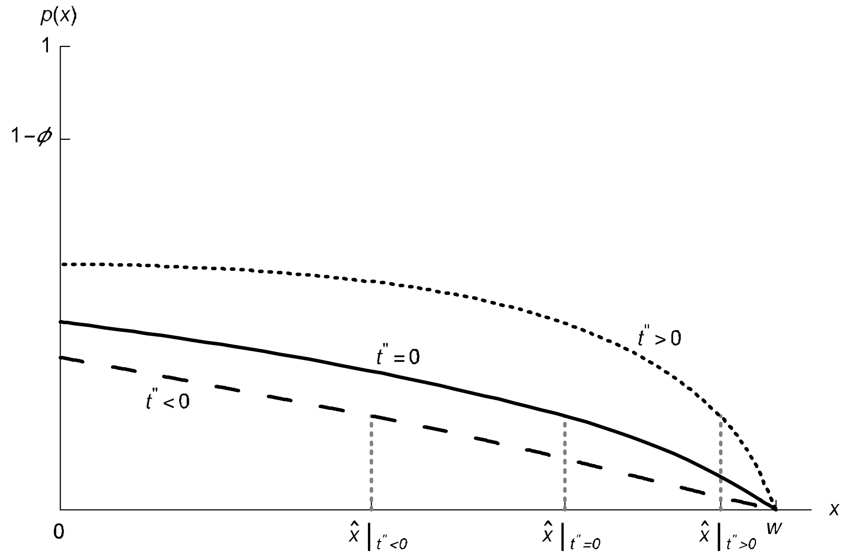


Figure 10: Audit function for a progressive, linear, and regressive tax function.

3.5.2 Risk Aversion

So far we have restricted the utility function to be linear. More generally, however, much evidence points towards risk aversion, which implies a utility function satisfying $U'' < 0$. Figure 11 illustrates $p(x)$ when taxpayers are risk neutral ($U(x) = x$) and when they are risk averse ($U(x) = x^{2/3}$). The audit function under risk aversion is seen to lie everywhere below the equivalent function when taxpayers are risk neutral. To understand this finding we apply Jensen's inequality to the condition $R\,(A^*, E^*) = 0$ to obtain

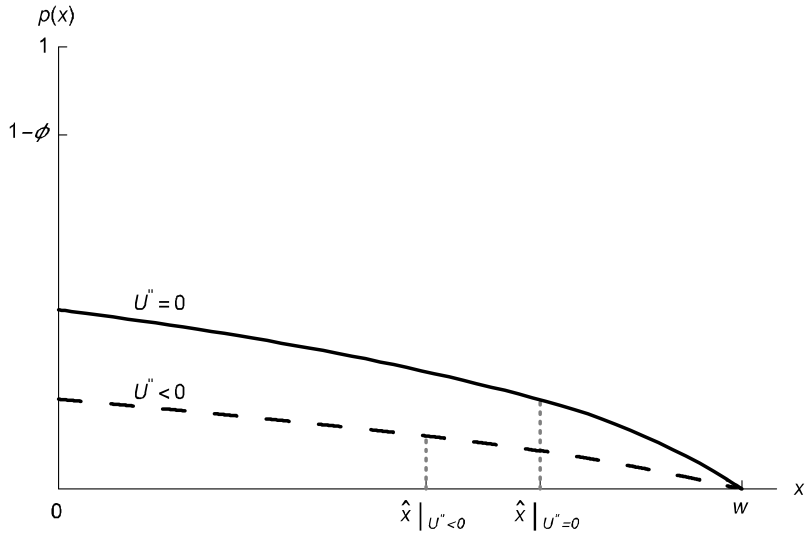


Figure 11: Effect of risk aversion.

$$p(x)\,U^a + [1 - p(x)]\,U^n = U^h \leq U\left(p(x)\,[w^a - S] + [1 - p(x)]\,w^n\right)$$

This inequality implies that $w^h \leq p(x)\,[w^a - S] + [1 - p(x)]\,w^n$, which is equivalent to $p(x) \leq \left[w^n - w^h\right] / \left[w^n - w^a + S\right]$. Under risk neutrality this inequality binds, so $p(x)$ must necessarily lie below the risk neutral level when risk aversion is introduced.

Furthermore, the audit function under risk neutrality is steeper than under risk aversion. Under risk neutrality an increase in declared income affects the taxpayer's payoff by the difference between the expected marginal return from truthful declaration and the expected marginal return of the lottery associated with under-declaration. However, if the taxpayer is risk averse, the expected marginal utility of an increase of x will also factor (positively) the reduction of risk. Hence, $p(x)$ in the risk aversion case is less sensitive to increases in the amount declared.

3.5.3 Variable Social Stigma

We now relax the previous assumption of a constant utility cost of social stigma by allowing for this cost to contain a variable component. We write

$$S(w-x) = \begin{cases} 0 & \text{if } x = w; \\ s + \psi[w-x] > 0 & \text{otherwise.} \end{cases}$$

where $\psi \geq 0$. When $\psi = 0$ we recover the specification of $S(\cdot)$ used in the previous section. Figure 12 compares the audit function in the two cases: one with a constant social stigma ($s = 3, \psi = 0$) and one with variable stigma ($s = 3, \psi = 0.9$). As can be seen from the Figure, the increase in ψ causes $p(x)$ to shift downward and become flatter. While the first effect is due to the absolute increase of the stigma cost, the second one is caused by variation in the marginal stigma cost. Indeed, for a unitary increase of declared income x, the taxpayer reduces his stigma by an amount ψ, hence, the reduction in $p(x)$ following an increase x is smaller the higher is ψ. In this way, holding the level of stigma constant, stiffer deterrence is needed when the stigma cost is dependent on evaded liabilities so as to counteract the stigma-relieving effect of an increase in the declaration.

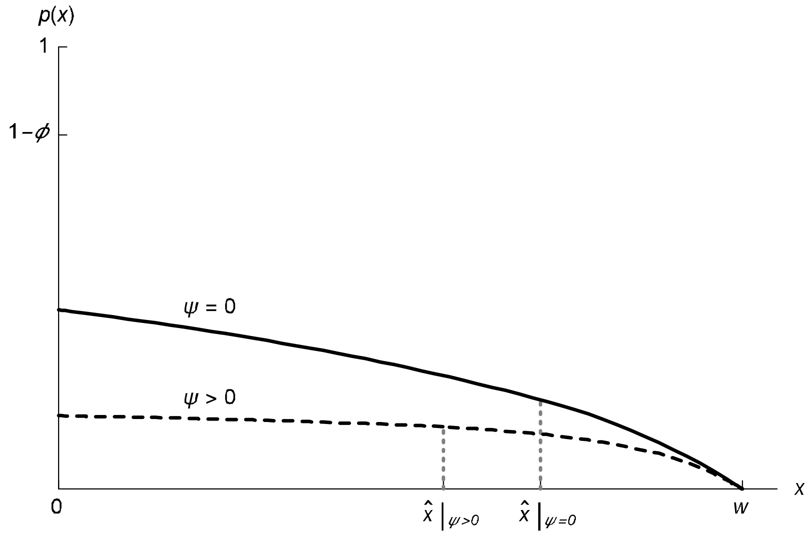


Figure 12: Effect of a variable component to social stigma.

3.6 Conclusion

In this chapter we investigated how accounting for the ability of individuals to avoid tax, as well as to evade tax, alters the conclusions for optimal auditing of models in which only tax evasion is possible. The nature of the avoidance activity we consider is not explicitly prohibited by law, but is unacceptable to the tax authority. Accordingly, if the tax authority learns of the avoidance, it moves (successfully in our model) to outlaw it ex-post.

Some key features of the literature that considers only evasion are preserved: we find that the audit function is a non-increasing function of declared income and, as in Chander and Wilde (1998), lower enforcement is required to enforce a regressive tax than to enforce a progressive tax. The model does, however, also yield new insights, in particular around the relationship between tax compliance and marginal tax rates. The evasion-only literature has encountered the so-called ”Yitzhaki puzzle” whereby stiffer marginal tax rates decrease incentives to be non-compliant. In our framework, however, the opposite applies: incentives to be non-compliant increase with marginal tax rates. The key to this result is that the incentives to avoid tax unambiguously increase following a raise of marginal tax rates. Thus, even though the incentives for evasion may worsen, nonetheless, the tax system becomes more costly to enforce, and overall compliance falls unless enforcement is stiffened. The theoretical result of a direct relation between tax rates and both avoidance and evasion matches the finding of Kleven et al. (2011). However, empirical investigations considering the relation between tax rates and non-compliance present mixed results. Further evidence of a positive link between the two has been reported empirically by Chiarini et al. (2013) while a negative one is presented in Slemrod (1985) and Feinstein (1991). Notwithstanding, we would like to stress that the field experiment of Kleven et al. 2011 benefits from a “designed” exogenous variation in tax rates that greatly improves the reliability of their inquiry. It would be of interest to compare our results with experimental evidence, unfortunately, to our knowledge there are no studies whose setting allow for both evasion and avoidance to be investigated.

A number of policy implication may be derived from the model. However, it is necessary to stress that, following the discussion in 3.3, our main focus is the audit function and other variables are considered as given in the analysis (i.e., may vary only in the "long-run").

Indeed we are able to understand further questions such as "which taxpayers are the most difficult (expensive) to make compliant?"; and "should tax auditing be geared to preventing avoidance or evasion?" On the first question, we find that in plausible circumstances it is the wealthiest taxpayer who is the most difficult to make compliant. While we know of no direct empirical evidence on this matter, our result chimes with the findings of attitudinal research regarding perceptions of the compliance of the rich (e.g., Wallschutzky 1984; Citrin 1979). The answer to the second question depends critically on (i) the level and shape of the penalties for evasion; and (ii) the competitiveness of the market for avoidance schemes (for this determines the share of the possible proceeds from avoidance that must be paid as a fee). If the penalty function is linear or concave then, irrespective of the tax function, a non-compliant taxpayer will engage purely in avoidance, or purely in evasion. Thus enforcement is focused entirely on one form of non-compliance or the other. When, however, the penalty function is convex (which seems quite likely empirically, given that smaller cases of tax evasion are typically punished through fines, but larger cases are punished through prison sentences) a non-compliant taxpayer may simultaneously want to avoid and evade tax, so enforcement must reflect both of these possibilities. We have shown that a taxpayer's preferred mix of avoidance and evasion moves in favour of avoidance as reported income decreases, as the competitiveness of the market for avoidance schemes increases, and as the social stigma associated with tax non-compliance falls.

Furthermore, the model provides useful insights on the effect that some "long run" parameters have on revenue collection costs. Fiscal variables directly affect collection efficiency: lower marginal tax rates and a reduced progressiveness of the tax schedule decrease the cost of revenue collection. A stiffer enforcement of tax laws imposing (relatively) high fines would allow for a reduction of enforcement costs. Finally, specific characteristics of taxpayers' population, i.e., a marked risk aversion or a high level of social stigma, would improve collection efficiency. However, it is necessary to note that enforcement of tax laws represents only one of a government targets. Indeed, the first part of the "long run" variables here considered plays a major role with respect to other government goals (economic policy, redistribution, enforcement, etc.) while the latest one would be barely affected by government intervention. Hence, the derivation of normative implication from the insights provided by the model must account for its partial-equilibrium nature and should be performed with caution.

We finish with some avenues for future research. First, it would be of interest to allow for imperfect audit effectiveness, as in Rablen (2014), Snow and Warren (2005a,b), for it might be that evasion and avoidance differ in the amount of tax inspector time required to detect them. Second, it might also be of interest to model more carefully the market for avoidance. In practice there are a range of providers of tax advice, ranging from those that offer solely tax planning, to those that are willing to offer aggressive (or even criminal) methods, making it important to understand the separate supply- and demand-side effects. As noted before, the literature does not provide any experimental study considering both avoidance and evasion. Further research along these lines would improve our understanding of non-compliance decision providing valuable insights to enhance deterrence policies. A last suggestion is to explore the effects of different forms of avoidance. We assume that avoidance permits some amount of income to be hidden from the tax authority, but an alternative modelling approach might be to assume that it allows some amount of income to be taxed at a lower rate.

3.7 Appendix

Proof of Proposition 6: The proof follows immediately from (4), (5) and (6).

Proof of Proposition 7: Differentiating in (7) we obtain that, if $A^* = 0$, then

$$\begin{aligned}
\frac{\partial p(x)}{\partial \overline{w}} &= \frac{t'(\overline{w})\left[1 - p\left(x\right)\right] f'}{f + s} > 0; \\
\frac{\partial p(x)}{\partial x} &= -\frac{\left[1 - p\left(x\right) f'\right] t'(\overline{w})}{f + s} < 0; \\
\frac{\partial p(x)}{\partial s} &= -\frac{p\left(x\right)}{f + s} < 0; \\
\frac{\partial p(x; \varepsilon f)}{\partial \varepsilon} &= -\frac{p\left(x\right) f}{f + s} < 0; \\
\frac{\partial p(x, \varepsilon t)}{\partial \varepsilon} &= p\left(x\right)\left[1 - p\left(x\right) f'\right] > 0; \\
\frac{\partial p(x)}{\partial \phi} &= 0.
\end{aligned}$$

The comparative statics when $A^* = w - x$ follow similarly.

Proof of Lemma 1: We first prove $p\left(x\right) f' < 1$ From (14), (16) and (18), if there exists a $\hat{w} \leq \overline{w}$ such that $p(x; A, w)$ attains the value $p(x; A, \hat{w}) = \left[f'\left(t\left(\hat{w}\right) - t\left(x + A\right)\right)\right]^{-1}$ then $p(x; A, \hat{w}) = \max_w p(x; A, w)$ – for if (13) defines a maximum in A, as assumed, then (13) defines a maximum in w. $p(x; A, \hat{w})$ is maximised in A when $\hat{A} = 0$ (as $f'' > 0$ for there to be an interior A^*), so $\hat{w} \neq w^*$ for, by assumption, if it were that $\hat{w} = w^*$ then $p(x; A, \hat{w})$ would be maximised for an interior value of A. Hence we have $\left[f'\left(t\left(\hat{w}\right) - t\left(x + A^*\right)\right)\right]^{-1} > p(x; A^*, w^*) = p\left(x\right)$. As this will hold for every $\hat{w}$, we have $p\left(x\right) f' < 1$. If $\partial p(x; A, w)/\partial w > 0$ everywhere then there does not exist a $\hat{w} \leq \overline{w}$ such that $\partial p(x; A, w)/\partial w =$

0. We note that it cannot be that $\partial p(x;A,w)/\partial w < 0$ everywhere as $\partial p(x;A,w)/\partial w|_{A=w-x=0} = \phi t'(x) / [s + f(0)] > 0$. In this case $p(x;A,w)$ is maximised at $w = \overline{w}$ and satisfies $p(x;A,\overline{w}) < [f'(t(\overline{w}) - t(x+A))]^{-1}$. An analogous argument to that above then establishes that $p(x) f' < 1$. Then, from (15), we may set $p(x) = \phi [f' - 1]^{-1}$ in $p(x) f' < 1$ to obtain $[1-\phi] f' > 1$. That $p(x) < 1 - \phi$ follows immediately.

Proof of Proposition 8: The comparative statics of a pivot around $(z, f(z)) = (0,0)$ are found by writing $f(\cdot)$ as $\varepsilon f(\cdot)$, differentiating with respect to ε, and then examining the resulting derivative as $\varepsilon \to 1$. The pivot of the tax function is performed analogously. The comparative statics of a shift of the tax function are found by replacing $t(\cdot)$ with $t(\cdot) + \varepsilon$, differentiating with respect to ε, and then examining the resulting derivative as $\varepsilon \to 0$. When $A^* \in (0, w - x)$ we use the implicit function theorem in (13) to obtain:

$$
\begin{aligned}
\operatorname{sgn}\left(\frac{\partial A^*}{\partial s}\right) &= -\operatorname{sgn}\left(\phi t'(A+x)\right) < 0; \\
\operatorname{sgn}\left(\frac{\partial A^*}{\partial \phi}\right) &= -\operatorname{sgn}\left(f + [t(A+x) - t(x)]f' + s\right) < 0; \\
\operatorname{sgn}\left(\frac{\partial A^*}{\partial \overline{w}}\right) &= \operatorname{sgn}\left([1-\phi]f' - 1 + [w^n - w^h]f''\right) > 0; \\
\operatorname{sgn}\left(\frac{\partial A^*}{\partial x}\right) &= -\operatorname{sgn}(\{[1-\phi]f' - 1\}\, t'(x) + [w^n - w^h]t'(A+x)f'') < 0; \\
\operatorname{sgn}\left(\frac{\partial A^*(\varepsilon f)}{\partial \varepsilon}\right) &= \operatorname{sgn}(s\phi + t(\overline{w}) - t(x)) > 0; \\
\operatorname{sgn}\left(\frac{\partial A^*(\varepsilon t)}{\partial \varepsilon}\right) &= \operatorname{sgn}([t(\overline{w}) - t(A+x)][w^n - w^h]f'' + \\
&\qquad \{[1-\phi]f' - 1\}[t(\overline{w}) - t(x)] > 0;
\end{aligned}
$$

and when $w^* \in (x, x+A)$ we use the IFT in (14) to obtain

$$\begin{aligned}
\operatorname{sgn}\left(\frac{\partial w^*}{\partial s}\right) &= \operatorname{sgn}\left(t'(w^*)\right) > 0; \\
\operatorname{sgn}\left(\frac{\partial w^*(\varepsilon f)}{\partial \varepsilon}\right) &= -\operatorname{sgn}\left(\frac{st'(w^*)}{[f+s]^2}\right) < 0; \\
\operatorname{sgn}\left(\frac{\partial w^*}{\partial x}\right) &= \operatorname{sgn}\left(\frac{t'(x)}{t'(w^*)}\right) > 0; \\
\operatorname{sgn}\left(\frac{\partial w^*}{\partial \overline{w}}\right) &= \operatorname{sgn}\left(0\right) = 0 \\
\operatorname{sgn}\left(\frac{\partial w^*(\varepsilon t)}{\partial \varepsilon}\right) &= -\operatorname{sgn}\left(\frac{f''t'(w^*)}{\left[f'\right]^2}\right) < 0; \\
\operatorname{sgn}\left(\frac{\partial w^*}{\partial \phi}\right) &= \operatorname{sgn}(0) = 0.
\end{aligned}$$

Turning to $p\,(x)$, we use the IFT in (2) along with (13) or (14) to obtain

$$\begin{aligned}
\operatorname{sgn}\left(\frac{\partial p(x)}{\partial s}\right) &= -\operatorname{sgn}(w^n - w^h) < 0; \\
\operatorname{sgn}\left(\frac{\partial p(x;\varepsilon f)}{\partial \varepsilon}\right) &= -\operatorname{sgn}([w^n - w^h]f) < 0; \\
\operatorname{sgn}\left(\frac{\partial p(x)}{\partial x}\right) &= \begin{cases} -\operatorname{sgn}\left(\frac{1-p(x)-\phi}{w^n-w^c+s}\right) < 0 & \text{if } A^* \in (0, w-x)\,; \\ \operatorname{sgn}(0) = 0 & \text{if } w^* \in (x, x+A)\,; \end{cases} \\
\operatorname{sgn}\left(\frac{\partial p(x)}{\partial \overline{w}}\right) &= \begin{cases} \operatorname{sgn}(\{[1-\phi]f' - 1\}) > 0 & \text{if } A^* \in (0, w-x)\,; \\ \operatorname{sgn}(0) = 0 & \text{if } w^* \in (x, x+A)\,; \end{cases} \\
\operatorname{sgn}\left(\frac{\partial p(x;\varepsilon t)}{\partial \varepsilon}\right) &= \begin{cases} \operatorname{sgn}([1-\phi]f' - 1) > 0 & \text{if } A^* \in (0, w-x)\,; \\ \operatorname{sgn}(0) = 0 & \text{if } w^* \in (x, x+A)\,; \end{cases} \\
\operatorname{sgn}\left(\frac{\partial p(x)}{\partial \phi}\right) &= \begin{cases} -\operatorname{sgn}\left([1-\phi]f' - 1\right) < 0 & \text{if } A^* \in (0, w-x)\,; \\ \operatorname{sgn}(0) = 0 & \text{if } w^* \in (x, x+A)\,; \end{cases}
\end{aligned}$$

Chapter 4

Italian municipalities efficiency: A conditional frontier model approach

4.1 Abstract

In most developed countries local governments are responsible for a relevant share of the public services. However, the assessment of local government efficiency is a complex task that must account for the heterogeneity of local conditions. Hence, the estimation of relative efficiencies must factor in for the impact of contextual and exogenous variables on the production process. In this paper, we develop a composite local government output indicator with an unprecedented coverage of the productive activities performed by the major Italian municipalities. Using the output indicator along with balance sheet data of the expenses as an input, we apply conditional frontier models to provide an assessment of

municipalities efficiency scores. In the first stage, we evaluate the degree of improvements attainable by any municipality with respect to the *relevant* "best-practice" frontier. In the second stage, we investigate the role of non-discretionary inputs on performance.

4.2 Introduction

Local governments, especially in developed countries, provide for a sizable part of the public good and services supply. The considerable amount of public spending and decision making attributed to local governments calls for an evaluation of their performances. Indeed, efficiency assessment of local governments may provide meaningful information to numerous agents, first among others, the central government, which is interested in the design of performance-based transfers and constraints for the local level. In fact, the decentralization of both taxation and provision determines different effective tax bases and expenditure needs depending on potential spill-overs (Oates 1999) and local socio-economic dimensions. Hence, federal taxes are usually integrated with transfers from the central government based on fiscal efficiency and fiscal equity considerations (on the topic, see Prud'homme 1995; Oates 2005 and Shah 2012). Measures of performances may then be used, along with the socio-economic composition and past expenditures, to determine transfers and constraints imposed on local branches (Oates 1999; Bordignon 2001; Lockwood and Porcelli 2013).

Acknowledgements: We are grateful to Jilles Saint-Paul and Andrea Vindigni for insights and helpful comments on an earlier version of the manuscript. We thank Simone Faggi and Ivan Marzocco whose first-hand experience greatly improved our understanding of municipality structure. Any errors are our own and those we thank may not agree with the interpretations provided in this paper.

Moreover, local politicians may take advantage of efficiency assessments as a relevant information on the production function of the local bureaucratic apparatus. This information may be used to better evaluate the electoral returns offered by different options to increase, therefore, their political power (see Brennan and Buchanan 1980; Besley and Case 1995; Persson et al. 1997; Besley and Coate 2003) and/or to implement policies to attract mobile factors of production (on the determinants of firm localization, see Papke 1991; Patel and Vega 1999; Buettner and Ruf 2007).

Finally, efficiency assessment would provide voters with a signal of the managerial skills of the local incumbent (see the seminal contribution by Seabright 1995 and, more recently, Besley and Smart 2007). While in the literature it is disputed whether additional information would result in positive or negative net welfare effects (see Ferejohn 1986; Baron 1994; Dewatripont et al. 1999; Besley and Smart 2007), it is apparent that the electorate, or at least a part of it, is interested in the matter.

Evaluation of *allocative* efficiency may be performed using both input and output quantitative indicators along with unit prices as weights. However, in the municipalities case, market prices for outputs are usually not available. To overcome this issue, literature proposes to estimate the production function frontier (parametrically or non-parametrically) and to derive *technical* efficiency scores on the basis of relative distances of inefficient observations from the frontier (see e.g. Farrell 1957; Dorfman et al. 1958; Charnes et al. 1979). Although the use of parametric estimation would enhance estimation efficiency, we adopt the non-parametric approach due to the lack of a convincing theory in the literature on the true form of the production function (or of the distribution of inefficiencies) for municipalities. The non-parametric framework offers two main models, the Data Envelopment Analysis (DEA) and the Free Disposal Hull (FDH), where the main differences between the two approaches are related to the fact that the former requires *i)* the assumption on the convexity of technology to hold, and *ii)* presents a better rate of convergence (see the seminal work of Deprins et al. 1984 and, among the latest contributions, Kneip et al. 2011). In Daraio and Simar (2007b) it is provided a test to perform a data-driven choice between the two models.

The heterogeneity of exogenous conditions faced by the unit of interest has a deep effect on efficiency evaluation and, if not considered in the estimation, leads to biased and unfair efficiency measures: some units score poorly because they are impaired by exogenous factors, while other result efficient without merit. Firms may deal with the exogenous conditions by changing their location in search of the most "suitable" one for the production process (Kravis and Lipsey 1982; Wheeler and Mody 1992; Flores and Aguilera 2007). Conversely, municipalities cannot exploit the relocation option and the effect of external variables is arguably even more pronounced. The methodologies proposed to account for non-discretionary variables in the non-parametric setting usually entail multi-stage models. In the first step an initial measure of efficiency is computed, while in the second step the effects of exogenous factors are estimated and washed out from the initial efficiency scores (see among others: Loikkanen et al. 2005; Balaguer-Coll and Prior 2009) by means of different econometric techniques, usually OLS, censored regression or Tobit models. However, Simar and Wilson (2011b) show the inconsistency of many of these methodologies and in Simar and Wilson (2007) a theoretically sounding model, which entails a bootstrapped-

DEA coupled with Tobit regression, is proposed. Nevertheless, the last-mentioned method requires the separability condition to hold, namely: "the support of the output variables does not depend on environmental variables" (Simar and Wilson 2011b, p. 207) and hence, neither the shape nor the level of the boundary of the attainable set should vary with the exogenous factors. Whether the separability condition holds or not can be assessed by means of the test proposed in Daraio et al. (2010), but it has to be noted that its requirement is rather strong. This restriction has been relaxed thanks to a methodology presented in Daraio and Simar (2005) and Daraio and Simar (2007b) that only requires the partial separability condition to hold, namely: "the shape of the boundaries production frontier would not change with the level of output" (Badin et al. 2012, p. 821). This approach introduces the exogenous factors directly in the first step, i.e. in the bootstrapped-conditional-DEA (or bootstrapped-conditional-FDH), and then performs a second step by means of a smoothed regression to evaluate the mean/variance effects of exogenous variables on the efficiency scores.

The non-parametric frontier estimation has been applied to a wide rage of topics (for a thorough survey see Emrouznejad et al. 2008 and the references therein). Among the literature on municipality efficiency, one strand is focused on a particular local service e.g., fire protection (Kristensen 1983; Bouckaert 1992), garbage collection services and solid waste management (Bosch et al. 2000; Worthington and Dollery 2001; Sarkis and Dijkshoorn 2007; García-Sánchez 2008; Huang et al. 2011; Rogge and De Jaeger 2012), general administration and register offices (Kalseth and Rattsø 1995; Settimi et al. 2014), infrastructure and equipment (Prieto and ZofIo 2001), police services (Carrington et al. 1997; Diez-Ticio and

Mancebon 2002; Garcia-Sanchez 2009; Verschelde and Rogge 2012), public library services (De Witte and Geys 2011), urban public transportation (Pina and Torres 2001; Boame 2004), or water management (Picazo-Tadeo et al. 2009; Gupta et al. 2012). Conversely, another strand evaluates global performances, i.e. it considers a more comprehensive list of goods and services provided by local governments; this literature includes application to numerous countries, such as: Australia (Worthington 2000), Belgium (De Borger and Kerstens 1996; Geys 2006; Ashworth et al. 2014), Brazil (De Sousa and Stošić 2005), Finland (Loikkanen et al. 2005), Germany (Geys et al. 2010; Kalb et al. 2012), Greece (Athanassopoulos and Triantis 1998; Doumpos and Cohen 2014), Italy (Boetti et al. 2012; lo Storto 2016), Japan (Nijkamp and Suzuki 2009), Portugal (Afonso and Fernandes 2006, 2008, Da Cruz and Marques 2014), and Spain (Balaguer-Coll et al. 2007; Benito et al. 2007; Gimnez and Prior 2007; Balaguer-Coll and Prior 2009; Balaguer-Coll et al. 2010; Benito et al. 2010; Cordero and Pedraja-chaparro 2016), and USA (Grossman et al. 1999).

Given data availability and the difficulties arising from the imputation of general inputs to particular services, the present study considers the whole set of functions undertook by Italian municipalities to provide an appraisal of local government efficiency. In the scientific literature, the closest analysis to the present one are Boetti et al. (2012) and lo Storto (2016). While the latter contribution computes an efficiency assessments for (almost) the same set of municipalities here considered, the former one is focused on Piemonte's municipalities only. These studies employ Stochastic frontier Analysis and a two-stage DEA-Tobit as methodology with lo Storto 2016 also providing confidence intervals for DEA estimates by means of bootstrapping (Simar and Wilson 2007).

While far from being complete, the literature review here proposed highlights the importance given by the scientific community to the non-parametric measure of efficiency for the evaluation of local government performance. The present work implements state-of-the-art methodologies in the field (an organic presentation is provided in Bădin et al. 2014) to compute an efficiency assessment for the major Italian municipalities. To this end, a data set with an unprecedented coverage of local government interventions is developed along with a rich set of variables to account for the exogenous conditions. Three different external/exogenous dimensions are investigated. To provide a sound representation of the applied methodology we consider geographical conditions, i.e. we study whether physical constraints impact on the level of attainable efficiency. We introduce an adjusted measure of altitude to account for this dimension, that is, for each municipality we consider the joint effects of the level of altitude and its variability in the municipality territory. Therefore, we test whether being in a territory characterized by a high level of altitude and/or of dispersion (i.e. a high mix of mountains, hills and level ground) may impact as an unfavourable input for the provision of local public goods and services. We also introduce a measure of social environment, i.e. we investigate whether efficiency scores are influenced by the level of crime in the municipality. To this end, we take into account the average number of crimes reported per thousand residents in a three years timespan before 2011. The characteristics of the resident population depend on many socio-economic drivers and both local and central policies affect migration net-flows and the local social environment alike. However, we argue that this is a viscous process, which we assume requires at least a couple of years to modify the local social conditions. We expect that the sign of this relationship is negative, i.e. being in an unfavourable social environment (more crimes per inhabitant) negatively affects the provision by municipalities of an efficient mix of inputs and outputs. Finally, we opt for a measure of indebtedness to gauge the potential effects of financial constraints. We use the average values of the stock of municipality debt over an interval of four year prior to the computation of the efficiency scores to investigate the effect of per-capita local debt on efficiency frontier. One of the major dangers in a decentralised fiscal system is the presence of soft budget constraints (SBC), i.e., the inability of the central government to credibly rule out the possibility of bailout in case of local default. In presence of SBC, local branches of government have an incentive to increase their debt above the optimal level leading to suboptimal policies and increasing bankruptcy risk. In order to manage this danger, the design of fiscal relations between the multiple levels of government becomes crucial (Oates 2005). Financial constraints imposed by the central government on the local levels has been debated extensively in the public finance field and in the political economy literature (i.e. Persson and Tabellini 1996; Rodden 2002). Especially relevant to the present study is the EU Stability and Growth Pact (SGP) (see EC Commission 1990; EC Commission: The Commission Services 1995; EC Commission 1996) that imposes constraints on the level of debt. Academic debates on the optimality of SGP have ensued both at international (Beetsma and Uhlig 1999; Buti et al. 2003; Feldstein 2005) and at local levels (for the Italian case, see Balassone and Zotteri 2001; Brugnano and Rapallini 2009; Chiades and Mengotto 2015). We guess that the estimation of the impact of the historical stock of debt on local efficiency, along with the assessment of managerial efficiencies accounting for the level of local debt, may contribute to this debate.

Aim of the present work is to contribute to this literature with an investigation that exploits one of the most recent developments in the non-parametric efficiency literature, namely, robust-conditional efficiency measures. Our findings indicate that municipalities characterized by a territory with an high altitude level and/or dispersion are hindered in their provision activity. Thus, a fair rank of efficiency scores should consider the location of municipality and the features of its territory. Similarly, the social context seems to matter and negative social conditions are, as expected, identified as an unfavourable input. Finally, the level of indebtedness appears to be marginal, having no visible effects on the likelihood to reach the efficient frontier. These unidimensional analysis illustrate some of the diverse perspective that can be adopted while measuring "pure" managerial efficiency, i.e., measures where the effect of external/exogenous variable is removed by means of a two-stage procedure. We also provide a multi-dimensional study which combines both physical and social measures. This enriched view mostly confirms the univariate results and allows to identify conjoint effects neglected otherwise besides providing a more thorough removal of exogenous effects from managerial efficiencies.

The paper is organized as follows: Section 4.3 describes the data set and the corresponding sources; Section 4.4 presents the methodology in a step-by-step procedure to highlight the construction of both conditional and unconditional efficiency scores and to introduce the two-stage analysis; Section 4.5 shows the results on local efficiency scores and performs unidimensional and multidimensional second-stage analysis; Section 4.7 concludes and presents avenues for furher research.

4.3 Data Description

Our study focuses on performance evaluation at year 2011. The reason behind this choice relies on both data quality and coverage of the variables used to assess efficiency levels. Year 2011 represents a valuable point in time to describe the status of Italian society since it corresponds to the latest census made by the Italian National Institute of Statistics (ISTAT 2011).

Our units of interest (i.e. DMUs, Decision Making Units) are Italian administrative centers. There are about 8,000 municipalities in Italy. We select cities corresponding to the province capitals, hence the analysis covers the most important municipalities in terms of both populations and territory sizes. Among the 110 province capitals, 10 of them correspond to unions of municipalities[1]. We prefer to discard these cases due to data coverage and comparability issues with the other administrative centers[2].

The analysis presents a static ranking of Italian municipality[3] efficiencies. Usually, the assessment of municipal efficiency levels is measured relating local expenditures, as a whole or partitioned into budget items, to a list of output targets (Worthington and Dollery 2000). Our framework includes a list of inputs which stand for the amount of money (in Euro) each DMU spends for its main functions. Inputs are retrieved from the Italian Ministry of Internal Affairs[4]. For each municipality, we use

[1]These province capitals are: Barletta-Andria-Trani, Carbonia-Iglesias, Forlí-Cesena, Massa-Carrara, Medio Campidano, Monza e della Brianza, Ogliastra, Olbia-Tempio, Pesaro-Urbino, Verbano-Cusio-Ossola.

[2]For the same reasons, we drop Aosta due to the high number of missing values affecting it.

[3]For simplicity, hereinafter we refer to the DMUs as municipalities.

[4]Data is available at the link: `http://finanzalocale.interno.it/apps/floc.`

financial statement items to compute the total expenditures for every function. The list of functions that we include is[5]: i) general activities related to administrative and control tasks; ii) local police; iii) education and schooling; iv) infrastructures and transportation; v) environment; vi) welfare. Hence, we consider the most relevant municipalities' functions for which corresponding outputs are available[6]. The construction of the outputs vector is based on several sources[7]: ISTAT, Istituto Tagliacarne, the Italian Ministry of Internal Affairs, the Italian Ministry of University and Research, and Legambiente. For each input we attempt to provide a parallel with the corresponding outputs. Due to the difficulty to identify the (single) most appropriate output for each input, we prefer to consider small lists of outputs for each input. Subsection 4.4.6 will explain how outputs are treated to be included in the computation of the efficiency scores. Below, Table 2 shows the list of inputs and the corresponding outputs, while Table 3 exhibits the descriptive statistics for each input and output.

Scholars study the impact on efficiency levels of a wide list of environmental or context indicators (see e.g. De Borger and Kerstens 1996;

php/in/cod/4.

[5]There is a repartition of competencies among Italian State, Regions, Provinces (nowadays cancelled by the Law n.56 of 7 April 2014, but present in 2011) and Municipalities. In September 2003 the Constitutional Law n.3 modified the Title V of Italian Constitution, determining also the revision of the "Testo Unico degli Enti locali" (approved by d.lgs. 18th August 2000, n.267), which represents the collection of regulations on local public entities.

[6]In addition, the omitted functions represent only a marginal part of municipality total expenditures. These functions are: justice, management of cultural heritage, sport and leisure, economic development, and provision of services like utilities. Other public bodies are typically much more involved in these fields, where municipalities are usually in charge of residual activities.

[7]ISTAT is the Italian Institute of Statistics. Istituto Tagliacarne belongs to the Italian statistics system and it is focused on data sets from the Chambers of Commerce. Legambiente is a leading environmental association in Italy.

Table 2: Inputs-Outputs Definitions. We show in the first column the selected inputs corresponding to the main municipality functions. Second column refers to the related outputs for each input, while third column stands for the respective source of the output (Inputs are all retrieved from Italian Ministry of Internal Affairs).

Inputs	Outputs	Outputs Sources
General activities related to administrative and control tasks	Total Population	ISTAT
Local Police	Number of Vehicles	Ministry Internal Affairs
	Number of Km Travelled	Ministry Internal Affairs
Education and Schooling	Number of Classes	Ministry of University and Research
	Number of Students	Ministry of University and Research
	Number of Meals	Ministry Internal Affairs
	Number of Applications for the Canteen Service	Ministry Internal Affairs
	Number of Satisfied Requests for the Canteen Service	Ministry Internal Affairs
Infrastructures and Transportation	Light Points	ISTAT
	Km of Lit Roads	Ministry Internal Affairs
	Km of Urban Roads	Ministry Internal Affairs
	Km of Extra-Urban Roads	Ministry Internal Affairs
	Users of Public Transport Service	ISTAT
	Number of Km Travelled by Public Transport	Legambiente
Environment	Power of Photovoltaic Solar Panels installed on Municipal Buildings	ISTAT
	Population Connected to Urban Wastewater	ISTAT
	Solid Waste Treated	ISTAT
Welfare	Home Help with Social Care Services: Disability	ISTAT
	Home Help with Social Care Services: Elderly	ISTAT
	Home Help with Social and Healthy Care Services: Disability	ISTAT
	Home Help with Social and Healthy Care Services: Elderly	ISTAT
	Cash Benefits for Delivery of Social and Healthy Care at Home: Disability	ISTAT
	Number of Day-Nursery	ISTAT
	Number of Available Seats in Day-Nursery	Ministry Internal Affairs
	Day-Nursery: Number of Children Aged 0-2 Years Attending Infancy Day-Care Services	ISTAT
	Number of Applications for the Day-Nursery Service	Ministry Internal Affairs
	Number of Satisfied Requests for the Day-Nursery Service	Ministry Internal Affairs
	Supplementary or Innovative Services for Infancy Day-Care	Ministry Internal Affairs

Table 3: Inputs-Outputs Descriptive Statistics. We report for each input and output the corresponding values of: minimum, median, mean and maximum. Inputs are in thousands Euro. Outputs units are expressed in parenthesis.

	min	median	mean	max
INPUTS (in thousands Euro):				
General activities related to administrative and control tasks	4983	22110	54730	986800
Local Police	990	4774	15050	392900
Education and Schooling	991	8406	20540	465100
Infrastructures and Transportation	556	8615	32180	812819
Environment	1525	17000	42110	928600
Welfare	2194	14180	37610	690700
OUTPUTS:				
Total Population (Number of residents)	21642	88812	168274	2617175
Number of Vehicles	0	33	63	843
Number of Km Travelled (Thousands)	0	261	5296	470000
Number of Classes	210	616	1039	12684
Number of Students	4386	13252	22601	277814
Number of Meals	0	248766	678256	14784256
Number of Applications for the Canteen Service	0	2699	63998	4108200
Number of Satisfied Requests for the Canteen Service	0	2699	63994	4108200
Number of Light Points	2000	13156	21204	198863
Km of Lit Roads	0	297	477	7600
Km of Urban Roads	30	231	557	19516
Km of Extra-Urban Roads	0	200	612	29516
Users of Public Transport Service (Thousands)	87	4360	39138	1382392
Number of Km Travelled by Public Transport (Thousands)	130	2254	6965	159648
Power of Photovoltaic Solar Panels installed on Municipal Buildings (Kw)	0	57	336	9003
Population Connected to Urban Wastewater Treatment (Number of residents)	237	77189	150656	2496000
Waste Collection (Tons)	10365	49778	99566	1697186
Home Help with Social Care Services: Disability (Number of users)	0	5016	11842	321913
Home Help with Social Care Services: Elderly (Number of users)	172	1073	2109	20937
Home Help with Social and Healthy Care Services: Disability (Number of users)	0	586	2935	49725
Home Help with Social and Healthy Care Services: Elderly (Number of users)	0	201	783	9596
Cash Benefits for Delivery of Social and Healthy Care at Home: Disability (Number of users)	0	2267	8541	346535
Number of Day-Nursery	0	6	17	339
Number of Available Seats in Day-Nursery	0	250	610	13284
Day-Nursery: Number of Children Aged 0-2 Years Attending Infancy Day-Care Services	89	1196	2418	34245
Number of Applications for the Day-Nursery Service	0	345	912	19016
Number of Satisfied Requests for the Day-Nursery Service	0	230	632	11562
Supplementary or Innovative Services for Infancy Day-Care (Number of users)	0	146	318	3907

Byrnes and Dollery 2002; Balaguer-Coll et al. 2007; Schmidt 2008; Kalb 2010; Ammons 2014; Doumpos and Cohen 2014; lo Storto 2016, among others), such as: the institutional framework (e.g. legal and regulatory issues, political issues as authority and power), the social-cultural environment (e.g. traditions, traits and behaviors of the population; demographic distribution), the legacy conditions (e.g. presence of financial grants, fiscal constraints and historical level of debt), physical and natural conditions (e.g. climate, topography, geology as either resources or constraints), the presence of economies of scale and economies of scope.

We condition municipalities' efficiency scores on a list of external variables representing three main domains of analysis: i) physical constraints, as represented by the mix effect of the altitude of the municipality and the dispersion of its territory; ii) social environment, as expressed by the level of crime; iii) budget constraints, as synthesized by the historical stock of debts. These external variables stand for dimensions over which municipalities are supposed to do not have discretionary power. Following notation in literature, we indicate this list of external/exogenous variables as the Z vector. Accounting for external conditions that may impact on the effective capacity of municipalities to reach better efficiency levels allows to isolate the pure municipalities' efficiencies. Indeed, conditional measures evaluate DMU's efficiency with respect to the relevant (effectively achievable) frontier. This preliminary step is then followed by a second stage analysis where the pure managerial efficiencies are investigated. Section 4.4 will present and discuss the assumptions behind the choice of the appropriate Z vector.

Here, we briefly anticipate that the assumption of partial separability regarding the efficient frontier is crucial to compute conditional efficiency scores. Our external variables aim to include three distinct perspectives that are supposed to affect efficiency. Altitude and its variants are common measures applied to differentiate municipalities based on environmental conditions; here the intuition is that providing services may be more difficult for municipalities located at higher altitudes or with a heterogeneous composition of its territory. We employ measures of the level and of the standard deviation of altitude elaborated by ISPRA[8] developed using the Zonal statistics algorithm on ISTAT data. For each municipality we multiply its altitude by the variance of municipality territory to take into account the presence of mountains, hills and level ground. We label this indicator $Compound\ Altitude$. Crime is introduced to consider the "quality" of the population; although one might argue that even municipality managers can influence the overall level of crime, we still add this dimension since we believe the amount of crimes to be mostly related to comprehensive causes that are affected only marginally by the role of municipality policies. Indeed, crime deterrence is hardly directly affected by local policies while it is the central government the main actor in this context[9]. We compute the indicator $Crime\ Incidence$ as the average along the interval 2008-2010 of per-capita crimes reported to the judicial authorities. Finally, we investigate whether the stock of debts is a constraint to the provision of public services. Obviously, there might be endogeneity related to the funding sources that we include in the input vector. For this reason, we averaged over the 4 years before 2011 in order to ameliorate endogeneity issues.

[8]Italian National Institute for Environmental Protection and Research.

[9]The National authority's powers are exercised by the Minister of Public Affairs, to which the Law n.121 of 1st April 1981 has assigned the responsibility for the enforcement of public order and security.

The corresponding indicator $Funding$ is then normalized by the number of residents. Summary statistics of the external/exogenous measures are presented in Table 4.

Table 4: External Variables. First column reports the external/exogenous measures, while following columns indicate the corresponding values of: minimum, median mean and maximum. Last column shows the source where the variable is retrieved. Funding is in euro per capita while Crime Incidence is relative to one thousand residents.

	min	median	mean	max	Source
Compound Altitude	1.42	12050	48860	515900	ISTAT
Crime Incidence	39.96	37.88	20.40	82.23	Ministry Internal Affairs, ISTAT
Funding	40.84	920.70	1018	3536	Ministry Internal Affairs, ISTAT

4.4 Methodology

In the following subsections it is illustrated the practical implementation of the efficiency measures' estimation. Each subsection represents a building block of our efficiency distribution analysis; therefore, for the sake of clarity they are presented separately. Subsection 4.4.1 introduces the estimation framework and explains how the production process is defined both in the input $\times$ output Cartesian space and in the extended space which includes the external/environmental variables. Subsection 4.4.2 presents the choice of the optimal bandwidth required to compute the *conditional* efficiency scores; then, subsection 4.4.3 focuses on a robust version of the frontiers helpful to investigate the impact of the environmental variables on the inefficiency distributions, while subsection 4.4.4 shows the local analysis of efficiency scores as a function of these external variables. Subsection 4.4.5 discusses how to include the external variables in the efficiency analysis of DMUs to assess the effective managerial capacities of municipality majors while subsection 4.4.6 explains how data is processed to build the vectors used to compute the efficiency measures. Finally subction 4.4.7 illustrates the procedure used to identify and remove outliers.

4.4.1 General Setup

We briefly introduce the notation of the production process[10]. Let $X \in \mathcal{R}^p_+$ the *p*-dimension input vector used to produce the *q*-dimension output vector $Y \in \mathcal{R}^q_+$. In addition, consider $Z \in\in \mathcal{R}^r$ as the *r*-dimension vector of external or environmental variables that may affect the production process. Denote by $\Psi = \{(x, y)$ can produce $y\}$ the marginal *unconditional* attainable set, and let $\Psi^z = \{(x, y)|Z = z$ can produce $y\}$ the *conditional* attainable set, where we can observe that $\Psi = \bigcup_{z \in \mathcal{Z}} \Psi^z$ so that, by construction, for all $z \in \mathcal{Z}$, $\Psi^z \subseteq \Psi$. The role of Z can be multiple, influencing the achievable values for the combinations (X, Y), potentially affecting the shape of the distribution of the attainable set, or it may impact on the distribution of the inefficiencies without affecting the boundaries of the efficient frontier. Z may even shows mix effects, affecting both the shape of the frontier and the distribution of the inefficiencies, or, conversely, it can be also completely independent of (X, Y).

[10]For a deep analysis on efficiency methodologies and estimation techniques, the interested reader can refer to Daraio and Simar (2007a).

Daraio and Simar (2005) and Daraio and Simar (2007b) provide the general setup for the analysis of the joint contribution of the tern (X, Y, Z). Following previous work of Cazals et al. (2002), their formulation of the *conditional* process can be described as:

$$H(x, y|z) = Prob(X \leq x, Y \geq y|Z = z) \tag{1}$$

where $H(x, y|z)$ is the probability for a DMU operating at level (x, y) to be dominated by other DMUs facing the same environmental conditions z. This joint distribution can be decomposed as:

$$H_{X,Y|Z}(x, y|z) = F_{X|Y,Z}(x|y, z)S_{Y|Z}(y|z) \tag{2}$$

where $F_{X|Y,Z}(x|y, z) = Prob(X \leq x|Y \geq y, Z = z)$ and Equation 2 holds for all y such that the conditional survival function $S_{Y|Z}(y|z) = Prob(Y \geq y|Z = z) > 0$. Hence, the achievable couples (X, Y) given $Z = z$ determine the support of $H(x, y|z)$, that is Ψ^z. Similarly, the *unconditional* process can be represented as $H_{X,Y}(x, y) = Prob(X \leq x, Y \geq y) = F_{X|Y}(x|y)S_Y(y)$, where $F_{X|Y}(x|y)$ is the conditional distribution function of X and $S_Y(y)$ is the corresponding survival function of Y.

Literature proposes several measures to gauge the distance of a DMU operating at levels (x_0, y_0) to the efficient frontier (see for instance Debreu 1951; Färe et al. 2013; Farrell 1957; Shepherd 1970). In our setup, we rely on the Free Disposal Hull (FDH) estimator which is based on the free disposability assumption for Ψ and does not assume convex technologies as, indeed, required in the Data Envelopment Analysis (DEA). The FDH estimator of Ψ for the sample $\mathcal{X} = \{(X_i, Y_i), i = 1, ..., n\}$ is:

$$\hat{\Psi}_{FDH} = \{(x, y) \in \mathcal{R}_{+}^{p+q} | y \leq Y_i; x \geq X_i, (X_i, Y_i) \in \mathcal{X}\} \tag{3}$$

where the efficiency estimators are computed using a "plug-in principle", i.e. substituting the unknown quantities Ψ_{FDH} by their respective estimated values $\hat{\Psi}_{FDH}$. In our framework, we focus on how DMUs manage their inputs to determine certain level of outputs, hence we follow an input-oriented case, where the efficiency score of the DMU (x_0, y_0) is evaluated by means of the distance from this point to the estimated frontier of the input requirement set $\hat{C}(y) = \{x \in \mathcal{R}_{+}^{p} | (x, y) \in \hat{\Psi}_{FDH}\}$. Hence, the estimated *unconditional* input efficiency score for DMU (x_0, y_0) is defined as:

$$\hat{\theta}_{FDH}(x_0, y_0) = inf\{\theta > 0 | (\theta x_0, y_0) \in \hat{\Psi}_{FDH}\} = inf\{\theta | (F_{X|Y}(\theta x_0 | y_0) > 0\} \tag{4}$$

which can be estimated by solving the integer linear program

$$\begin{aligned} \hat{\theta}_{FDH}(x_0, y_0) = min\{\theta | y_0 \leq \sum_{i=1}^{n} \gamma_i Y_i; \\ \theta x_0 \geq \sum_{i=1}^{n} \gamma_i X_i, \sum_{i=1}^{n} \gamma_i = 1; \\ \gamma_i \in \{0, 1\}, i = 1, ..., n\} \end{aligned}$$

The nonparametric estimator in Equation 4 can be computed using the empirical version of the unknown conditional distribution of $F_{X|Y}(x|y)$, that is:

$$\hat{F}_{X|Y,n}(x|y) = \frac{\sum_{i=1}^{n} \mathbf{I}(x_i \leq x, y_i \geq y)}{\sum_{i=1}^{n} \mathbf{I}(y_i \geq y)} \tag{5}$$

where $\mathbf{I}(\cdot)$ is the indicator function which assumes value $\mathbf{I}(v) = 1$ if v is true and 0 otherwise. Due to the presence of the equality in the conditioning on Z in Equation 1, the *conditional* analog of Equation 4 requires some smoothing techniques. Daraio and Simar (2005) suggest the following kernel estimator:

$$\hat{F}_{X|Y,Z,n}(x|y,z) = \frac{\sum_{i=1}^{n} \mathbf{I}(x_i \leq x, y_i \geq y) K((z - z_i)/h)}{\sum_{i=1}^{n} \mathbf{I}(y_i \geq y) K((z - z_i)/h)} \tag{6}$$

where $K(\cdot)$ is an univariate kernel defined on a compact support and $h > 0$ is the corresponding bandwidth (Daraio and Simar 2005). Hence, the estimation of the *conditional* efficient scores can be interpreted as a restricted FDH program to points satisfying $||z_i - z|| \leq h \forall i \in [1, r]$.

4.4.2 Optimal Bandwidth Selection

To select the optimal bandwidth, we follow the approach proposed by Badin et al. (2010) who adapts and extends previous results in Hall et al. (2011) and Li and Racine (2008). Basically, we are looking for the bandwidth with the minimum Least Squares Cross Validation (LSCV) error for every point (x_i, y_i, z_i) $i \in [1, n]$ where n is the number of observations[11]. The set of bandwidths where the optimal one is searched for is generated in two stages. In the first one, unidimensional bandwidths (one for Y and one for every dimension of Z) are created by considering a grid of values. A desirable feature of the grid step is to ensure a similar "precision" along the different dimensions in order to achieve a balanced search in the mutidimensional space. Moreover, the biggest bandwidth considered should be wider than the range of the values assumed by the

[11]Conveniently, in Badin et al. (2010) a MATLAB routine for the task is provided. Furthermore, an equivalent code in R has been developed by us and available upon request

variable considered, thus including all the observations along the dimension considered, in order to exploit the discriminatory power of the LSCV procedure that: "[..] allows separating the influential from the irrelevant factors, by assigning to the irrelevant ones large smoothing parameters [..]" Badin et al. (2010, p. 639). Both features are implemented by multiplying the standard deviation of considered variable by a grid of values in $[0 + s, \frac{\max{(W)} - \min{(W)}}{\sigma(W)}]$ with step equal to s[12]. Finally, the second stage is performed by taking all the possible ordered selections with repetition of size $dim(Y) + dim(Z)$ among the unidimensional bandwidths thus obtaining the multidimensional ones.

4.4.3 Partial Frontier

The presence of potential outliers might affect the estimation of the efficiency scores. To provide a complementary study on the impact of Z on the efficiency distributions, we exploit partial frontiers (Daraio and Simar 2007a) which give measures of efficiency that are robust to outliers and extreme data points. In particular, we follow the order α-quantile frontiers introduced by Daouia and Simar (2007). For any $\alpha \in (0, 1]$, the order-α input efficiency score for a DMU operating at the level (x,y) is:

$$\theta_\alpha(x, y) = inf\{\theta | F_{X|Y}(\theta x|y) > 1 - \alpha\} \tag{7}$$

which converges to the Farrell-Debreu input efficiency score $\theta(x, y)$ when $\alpha \to 1$. Therefore, when $\theta_\alpha(x_0, y_0) = 1$ the corresponding DMU is efficient at the level $\alpha \times 100\%$, which means that this DMU is dominated by DMUs generating more output than y_0 with a probability $1 - \alpha$.

[12]The step used in the current work is $s = 0.1$

Conversely, when $\theta_\alpha(x_0, y_0) < 1$ the DMU (x_0, y_0) needs to decrease its input to the level $\theta_\alpha(x_0, y_0)x_0$ to reach the efficient frontier of $\alpha \times 100\%$. Finally, a DMU can be super-efficient with respect to the order-α[13] frontier when $\theta_\alpha(x_0, y_0) > 1$. The nonparametric estimator of the input efficiency score can be obtained using the plug-in procedure: $\hat{\theta}_{\alpha,n}(x, y) = inf\{\theta|\hat{F}_{X|Y}(\theta x|y) > 1 - \alpha\}$, whose statistical properties[14] have been presented in Daraio and Simar (2006).

4.4.4 Local Analysis

The environmental variables Z might influence the shape of the attainable frontier, thus *conditional* measures of efficiency appear more suitable to assess the level of effort a DMU must face to reach efficient levels. However, as pointed by Badin et al. (2010), DMUs should not be ranked simply by their *conditional* efficient scores since the impact of Z on the likelihood to reach the frontier may differ among different levels of Z. Conversely, a preliminary investigation on the local impact of Z on efficiency can be done independently of the correspondent level of inefficiency for each DMU. In fact, once the *conditional* and the *unconditional* efficiency scores have been computed, we can compare their ratio (R_I, for the input-oriented case) following the approach introduced by Badin et al. (2010):

[13] In the following, to facilitate readability and since no ambiguities are at stake, we will refer to the order α-quantile frontiers using indifferently the terms: alpha frontier, partial frontier and robust frontier.

[14] For finite sample, the order-α frontier does not include all the data points, being more robust to outliers than the standard envelopment estimators like FDH or DEA. In Daouia and Gijbels (2011a), Daouia and Gijbels (2011b) and Daraio and Simar (2007a) are reported some guidelines to select the appropriate order of the partial frontier to obtain robust estimates. The graphical analysis undertook to select the optimal alpha for the current work are reported in section 4.8, see Figures 28-29-30

$$R_I(x,y|z) = \frac{\theta(x,y|z)}{\theta(x,y)} = \frac{||x||\theta(x,y|z)}{||x||\theta(x,y)} = \frac{||x_y^{\partial,z}||}{||x_y^{\partial}||} \tag{8}$$

where $x_y^{\partial,z}$ and x_y^{∂} represent the projections of (x,y) on the *conditional* and *unconditional* efficient frontiers, respectively, along the x-axis and orthogonally to y. Hence, the ratio[15] R_I shows the shift of the frontier in the input orientation due to a certain impact of z and given a level y, regardless the inherent level of inefficiency for the DMU (x,y) (i.e. the modulus $||x||$).

In addition, the same study can involve partial frontiers (see subsection 4.4.3), providing therefore robust estimates of the distribution of the inefficiencies that complements the full frontier analysis. In particular, it is useful to analyze the impact of Z on partial frontiers because if $\Psi^z = \Psi$, then $\mathcal{R}_\alpha(z)$ can be used to study the local impact of Z on the shape of the distribution of inefficiencies over a grid of values of α like 0.99, 0.95, 0.90,..., 0.50, providing also a robust estimator of the boundary when α is close to 1. As pointed out by Bădin et al. (2014), the conditional average of R_I can be analyzed to detect whether locally Z has a significant impact on the boundary of the attainable set. Therefore, we define the corresponding values for the mean and the variance as:

$$\mu^z(P) = E[R(X,Y|Z=z)] \tag{9}$$

$$\sigma^{2,z}(P) = \mathcal{V}[R(X,Y|Z=z)] \tag{10}$$

[15]Consistent estimators of the ratios are obtained by plug-in procedures of the nonparametric estimators of the efficient scores. The rate of convergence is $\kappa = 4/((r+4)(p+q))$ for the full frontier case, while for partial frontier the rate is lower: $\gamma = 2/(r+4)$. See Daraio et al. (2010) for further details.

where P is an arbitrary DGP which generates (X, Y, Z). Nonparametric tools are exploited to estimate $\mu^z(P)$; below, we provide the framework for a univariate continuous Z, while for the multidimensional case see Bădin and Daraio (2011). Given a sample of n pairs $(Z_i, \hat{R}(X_i, Y_i|Z_i))$, for $i = 1, ..., n$, the conditional average $\mu^z(P)$ can be estimated as follows:

$$\hat{\mu}^z(P) = \sum_{i=1}^{n} W_n(Z_i, z, h_z)\hat{R}(X_i, Y_i|Z_i) \quad (11)$$

where we use local linear estimators for the weights $W_n(Z_i, z, h_z)$ (Fan and Gijbels 1996), while the localization is tuned by a bandwidth of appropriate size chosen using the LSCV criterion (Li and Racine 2007).

Finally, bootstrap procedures are applied to provide confidence intervals useful to make local inference on the impact of variables Z on the production process. We follow the methodology proposed by Bădin et al. (2014) to compute the bootstrap confidence intervals of $\hat{\tau}(z)$, since standard algorithm are not feasible because true $R(X_i, Y_i|Z_i)$ are unknown and pairs $(Z_i, \hat{R}(X_i, Y_i|Z_i))$ are not $i.i.d.$. Basically, we rely on the m out of n procedure introduced by Simar and Wilson (2011a) to obtain an approximation of the sampling distribution $\hat{\tau}_n^z - \tau^z(P)$. A step-by-step procedure for the confidence interval is[16]:

1 Given the sample $S_n = \{(X_i, Y_i, Z_i)|i = 1, ..., n\}$, estimate both the *unconditional* and the *conditional* input efficient scores, where for the latter the selection of the bandwidth reflects the LSCV criterion discussed in subsection 4.4.2, and determine the corresponding es-

[16]We rely on the formulation presented in the Appendix of Bădin et al. (2014) to highlight the key points of the bootstrap procedure. For an in-depth explanation, besides the references in this subsection see also Kneip et al. (2008), Kneip et al. (2011) and Politis et al. (2001).

timated ratios $\hat{R}(X_i, Y_i|Z_i)$.

2 Select the grid of values for Z, say $\{z_1, ..., z_k\}$ used to compute the nonparametric regression for $\hat{\tau}_n^{z_j}$ for $j = 1, ..., k$, where the bandwidth is chosen according to ASE.

3 For $m < n$ and large B (here $B = 2000$, n varies depending on the number of removed observations (due to missing data or ouliers) between $n = 83$ and $n = 93$[17], while m ranges from 50 to $n - 1$), repeat the following steps from 3.1 to 3.3 for each replication b in B.

 3.1 Draw a random sample $S^*_{m,b} = \{(X_i^{*,b}, Y_i^{*,b}, Z_i^{*,b})|i = 1, ..., m\}$ without replacement from S_n, maintaining the bandwidth selected in step 1 attached to the respective point $(X_i^{*,b}, Y_i^{*,b}, Z_i^{*,b})$.

 3.2 Determine the m ratios $\hat{R}^{*,b}(X_i^{*,b}, Y_i^{*,b}, Z_i^{*,b})$, for $i = 1, ..., n$ as in step 1. By doing so, rescale bandwidths obtained in step 1 multiplying them by the scale factor $(n/m)^{1/(r+4)}$.

 3.3 Using the same technique as in step 2, estimate the nonparametric regression over the points of the grid $\{z_1, ..., z_k\}$. Even in this case, bandwidths of step 2 are rescaled by the scale factor $(n/m)^{1/(r+4)}$.

4 For each point $j = 1, ..., k$, compute $\alpha/2$ and $1 - \alpha/2$ quantiles $(q^{*,z_j}_{m;\alpha/2}, q^{*,z_j}_{m;1-\alpha/2})$ of the B bootstrapped values $\hat{\tau}_m^{*,b,z_j} - \hat{\tau}_n^{b,z_j}$. Then, determine the corresponding confidence intervals of $\tau^{z_j}(P)$ at each point z_j as:

[17] $n = 93$ for $Z \in \{$ *Crime Incidence*, *Funding* $\}$, $n = 89$ if $Z \equiv$ *Compound Altitude*, $n = 83$ in the multidimensional investigation.

$$\tau^{z_j}(P) \in \left[\hat{\tau}^{z_j} - (m/n)^{2/(r+4)} q^{*,z_j}_{m;1-\alpha/2}, \hat{\tau}^{z_j} - (m/n)^{2/(r+4)} q^{*,z_j}_{m;\alpha/2}\right] \tag{12}$$

The selection of the optimal m for the subsampling is done also following Bădin et al. (2014). First, the steps three to four are replicated for $m \in [n/2, n-1]$ obtaining the relative confidence intervals. Second, according to Politis et al. (2001), we use the the average of the extremes of the intervals (for every point Z considered by the grid selected in point two) to evaluate a moving standard deviation[18] as a measure of the volatility associated with the considered m. We evaluated both the m-s that minimize the volatility for each observation and the one that minimizes the sum of volatility across all observation; the reported plots shows only the second one due to the reduced noisiness of the estimates.

4.4.5 Second Stage Analysis

Literature proposes several approaches to evaluate the impact of external variables on efficiency scores. Simar and Wilson (2007) and Simar and Wilson (2011b) discuss potential pitfalls that can be related to the inclusion of external variables in a second stage analysis framework, presenting some solutions to run a proper econometric analysis. In particular, we briefly remark that if the *separability*[19] condition does not hold, the

[18] the window that we selected is 15 given that the advised value of 3 resulted in noisy estimates.

[19] The *separability* condition refers to the input $\times$ output space with the inclusion of the Z space. It states that the enlarged set of attainable combinations is the Cartesian product $\Psi \times \mathcal{R}^r$, where r stands for the r-dimensional vector of Z external variables. Hence, this means that Z variables do not affect neither the attainable set Ψ nor the position of its frontier, but simply influence the distribution of the DMUs far from the frontier. For details see for instance Daraio and Simar (2007a).

interpretation of the *unconditional* efficient scores is misleading and the second stage analysis is therefore meaningless.

Here, we follow the approach introduced by Badin et al. (2010) where we basically study the average effect of z on $\theta(x, y|z)$. Obviously, efficient scores might be influenced by both y and z, but since the aim of this method is to detect the marginal effects of Z this justifies the analysis of $E[(\theta(X, Y|Z))|Z = z]$ as a function of z. As suggested by Badin et al. (2010), we adopt the following regression model:

$$\theta(X, Y|Z = z) = \mu(z) + \sigma(z)\epsilon \tag{13}$$

where $\mu(z) = E[(\theta(X, Y|Z))|Z = z]$, $\sigma^2 = \mathcal{V}[(\theta(X, Y|Z))|Z = z]$, $E[\epsilon|Z = z] = 0$ and $\mathcal{V}[\epsilon|Z = z] = 1$. $\mu(z)$ and $\sigma(z)$ are estimated nonparametrically: $\mu(z)$ by local constant methods (Pagan and Ullah 1999) which involve local smoothing techniques for the selection of the bandwidths with appropriate size (see Li and Racine 2007), while $\sigma^2(z)$ is estimated by regressing the squares of the residuals obtained for the estimation of $\mu(z)$ on Z (Fan and Gijbels 1996). Thus, $\mu(z)$ shows the average effect of z on the efficiency scores, while $\sigma(z)$ stands for the dispersion of the efficiencies as a function of z.

As a matter of fact, the resulting residuals appear informative to express the unexplained portion of the *conditional* efficiency scores:

$$\epsilon = \frac{\theta(X, Y|Z = z) - \mu(z)}{\sigma(z)} \tag{14}$$

In particular, if Z and ϵ do not present a strong correlation, ϵ represents the pure efficiency of DMU (x, y) since it is the cleaned off conditional efficiency where location and scale effect due to Z are removed

(Badin et al. 2010). Finally, we can use this quantity to evaluate the performance of municipality managers: large ϵ will refer to better performance, while small values of ϵ will represent poor managerial results.

However, the true input efficient scores are unknown and this implies we can exploit only the set of n estimators $\theta(X_i, Y_i|Z_i)$. This means that we have a sample of n pairs $(Z_i, \hat{\theta}(X_i, Y_i|Z_i))$, for $i = 1, ..., n$, which we use to estimate $\mu(z)$ and $\sigma(z)$. $\hat{\theta}(X_i, Y_i|Z_i)$ are estimators with n^{κ} rate of convergence, which imply that the estimator $\hat{\mu}_n(z)$ is a consistent estimator of $\mu(z)$ at a convergence rate equal to $n^{4/((r+4)(p+q))}$ (for details see Kneip et al. 2015).

4.4.6 Construction of the Input, Output and Environmental Vectors

In our setup, the set of attainable combinations (X, Y) is circumscribed to a simple representation in the Cartesian space where $p = 1$ and $q = 1$. In doing so, we need to find a methodology to collapse all the collected information for both inputs and outputs[20]. Inputs are expressed in monetary values, so we sum up the expenditures for each function to get the overall amount of money invested by each municipality to provide certain services.

The use of cost-related items to gauge municipalities performance is widely accepted in the empirical literature. Current expenditures are usually applied to obtain cost-efficient measures to assess municipality efficiencies (see e.g. Afonso and Fernandes 2008; Rogge and De Jaeger

[20]Literature proposes some approaches to reduce the multi-dimensionality issue related to the amount of inputs and outputs with respect to the sample size. The use of Principal Component Analysis is suggested e.g. in Daraio and Simar (2007a). The idea of composite indicators is also proposed in Afonso and Fernandes (2006) and Afonso and Fernandes (2008).

2013; Doumpos and Cohen 2014).

Conversely, outputs are expressed in different units and the amount of outputs that we collected differ between municipality functions. In addition, within each function these outputs reflect different aspects of the production process and we argue it would be unfair to select or discard some of them. Hence, for the outputs we follow the approach introduced by Afonso and Fernandes (2006) and Afonso and Fernandes (2008) and we provide a synthetic indicator which aggregates variables coherently to the correspondent municipality functions.

Since each variable has a particular dispersion, we normalize each value (for each DMU) by subtracting its mean value and then dividing this quantity by its standard deviation. Then, we sum up all these normalized values within the same municipality function. Finally, we divide by the effective number of variables which are available for that DMU within that municipality function. In this way we take into account also the event of missing values for some outputs. So, we end up with f normalized values for each DMU, where f is the number of municipality functions. This measure resembles the computation of the t-score, so that deviation from the mean value implicitly accounts for the dispersion of the variable considered. We adopt the label t-score hereinafter. In formula, the t-score of DMU i for a certain municipality function is:

$$\frac{\sum_{j=1}^{S}(X_{ij}-\mu_j)/\sigma_j}{S} \tag{15}$$

where S is the number of variables collected for DMU i for a certain municipality function, X_{ij} refers to the observation of DMU i for measure j, while μ_j and σ_j stand for the mean and the standard deviation of measure j, respectively. Therefore, each DMU will be characterized by a specific t-score for every function. To guarantee that non-negative values are considered in the FDH program, we simply shift t-scores corresponding to each municipality function by adding the absolute value of the respective minimum. Finally, we reduce output dimensionality by computing the mean values of the f t-scores for each DMU, so that any municipality function is assigned the same importance.

4.4.7 Outlier Detection

Given that the efficient frontier is estimated by extreme points of the sample, it follows that FDH measure (the same applies to DEA measure alike) is very sensitive to super-efficient outliers. As pointed out by Simar (1996), the identification and removal of outliers when using FDH methodology is crucial to the point that, may it be unworkable due to peculiar circumstances, other estimation methods should be preferred (such as stochastic frontiers). Building on the m-frontier developed in Cazals et al. (2002), Simar (2003) proposes a "Semi-automatic Warning Procedure" that may be used to detect outliers and Daraio and Simar (2007a) extended it to α-frontiers. The procedure is intended as a diagnostic tool to flag a set of observations that deserve closer inspection. For every sample point the procedure computes the number of observation dominating it and its "leave-one-out" partial frontier (α or m) measure of efficiency. The number of dominating units is relevant because it indicates the number of points used to estimate the efficiency measure and provides information on how close the point is to the edge of the support of the sample. Moreover, any sample point whose "leave-one-out" partial frontier measure is super-efficient (bigger than one in the input orientation or lower than one in the output one) when moving towards the "leave-one-out" full frontier (i.e., when α or m increases) stands out, in efficiency terms, from the cloud of sample points and is considered to be an extreme. Observations identified as extreme in both directions (input and output) and with a small number of dominating units, are therefore flagged potential outliers. Selection of an appropriate threshold to identify super-efficient observation may be performed by comparing, for different value of the threshold, the behaviour of the ratio of points above the threshold when the partial frontier converges to the full one. In absence of outliers, the curve of above-the-threshold observations should converge approximately linearly to the percentage of points having a leave-one-out FDH score greater than one (smaller than one for the output-oriented case). Conversely, strong deviation from linearity may be associated with the presence of outliers; a characteristic mark being the presence of an elbow effect consisting in sharp negative slope, followed by a smooth decreasing one in the graph. Given that an outlier could mask other outliers, the procedure of identification, verification and removal of outliers is performed iteratively.

Remarkably, accounting for exogenous variables in the efficiency measure affects (through the bandwidths) both the set of dominating units and the computed efficiency. Given a sample point, using conditional measures defines the conditional set of dominating units as a subset of the unconditional one whose points have distance, in the Z space, lower than the observation bandwidth. While the unconditional set of dominating units is informative with respect the position of an observation in the Y space (for the input orientation case, X in the output one) the conditional one provides analogous knowledge with respect to Z space.

Hence, in the present article, both measures have been considered while performing outlier identification. Outlier detection is undertook considering one exogenous variable at time leading to the identification of three partially overlapped sets. For $Compound\ Altitude$ the set is composed by: *Belluno, Biella, Bolzano, L'Aquila, Lecco, Pistoia, Sondrio, Reggio di Calabria, Rieti, Trento*; the variable $Crime\ Incidence$ identifies as outliers: *Catania, Milano, Napoli, Palermo, Roma, Torino* while for $Funding$ the observation of: *Catania, Genova, Milano, Palermo, Siena, Torino*, are removed. When performing multivariate analysis, the set of outliers considered is the union of the univariate outliers of the Z considered. For brevity we omit a part of the outliers identification procedure but we provide in section 4.8 the graphs of ratio of super-efficient observation vs α before and after the removal of outliers (see Figures 28-29-30).

4.5 Results

4.5.1 Local Analysis

Local analysis shows how the Z vector affects the production process of municipalities. We follow the methodology proposed by Badin et al. (2012), so the resulting local effects of each z on the shift of the frontier in the input-oriented framework can be analyzed by means of the full frontier ratio $\hat{R}_I(X_i, Y_i|Z_i)$. Partial frontiers ratios $\hat{R}_{I,\alpha}(X_i, Y_i|Z_i)$ will then be used to remove the effect of extreme data points that can hide the true efficiency levels. Finally, also the frontier with $\alpha = 0.50$ is analyzed to provide information on the effect of external measures on the middle of the distribution of inefficiency scores. To investigate the role of each z measure, we consider the marginal views of the conditional efficiency ratios on both levels of z and y, separately. We focus here on the impact of each z individually, while in Section 4.5.4 we will show the impact of the multi-dimensional case. In each local analysis, y stands for the $t-scores$ as introduced above in Section 4.4.6. Below, we summarize the local effects of each z:

- $z = Compound\ Altitude$. As shown in Figure 13, there is not a clear effect on the distributions of inefficiencies as a function of y. This is more visible for the robust version of the frontier ($\alpha = 0.98$). The investigation of Figure 13 reveals also the presence of a modest and increasing effect of z on the conditional efficiency scores for both the full and robust versions of the frontier. However, the median case shows a negative slope, suggesting the presence of some favourable effects on the frontier corresponding to the middle of the distribution. The support of the attainable set seems to slightly depend on $Compound\ Altitude$, but the probability of being far from the frontier (being less efficient) might be decreasing for municipalities with higher values of z. It is worth noticing that efficiency ratio seems not to be affected by y, meaning that there is not interaction between the effect of $Compound\ Altitude$ on the frontier and the levels of y. This motivates the marginal analysis which we will present in Sub-Sections 4.5.2 and 4.5.3.

- $z = Crime\ Incidence$. The relationships between z and conditional measures of efficiencies are positive in both the full and partial frontiers ($\alpha = 0.99$), thus supporting the interpretation that higher levels of $Crime\ Incidence$ act as an unfavourable input of the production process (Figure 14). This is also confirmed by the plot corresponding to $\alpha = 0.50$, which shows a slightly positive slope too, albeit of lower magnitude. This suggests that $Crime\ Incidence$ impacts on the shift of the frontier and less on the distribution of inefficiencies. Also in this case, y does not show a visible effect on the inefficiency ratio, meaning that there is not interaction between y and the effect of $Crime\ Incidence$ on the frontier. This supports the marginal analysis which we will discuss in Sub-Sections 4.5.2 and 4.5.3.

- $z = Funding$. The role of y on the distribution of conditional efficiencies is minimal as shown in Figure 15. Similarly, the effect of $Funding$ on conditional efficiencies appears to be absent as exhibited in both the full and the partial frontiers ($\alpha = 0.98$). However, the distribution of inefficiencies as a function of z at the middle of the distribution seems to exhibit a modest but positive relationship. This suggests that $Funding$ may affect distribution of inefficiencies and less the shift of the frontier. No interaction seems to be present between y and the ratio of efficiency when conditioning on $Funding$, supporting the marginal analysis which we will present in Sub-Sections 4.5.2 and 4.5.3.

Figure 13: Compound Altitude. For each block, the first plot refers the marginal impact of z on the efficiency ratios, while the second plot shows the marginal impact of y. Blocks stand for: full frontier, robust version ($\alpha = 0.98$) and middle distribution ($\alpha = 0.50$), respectively.

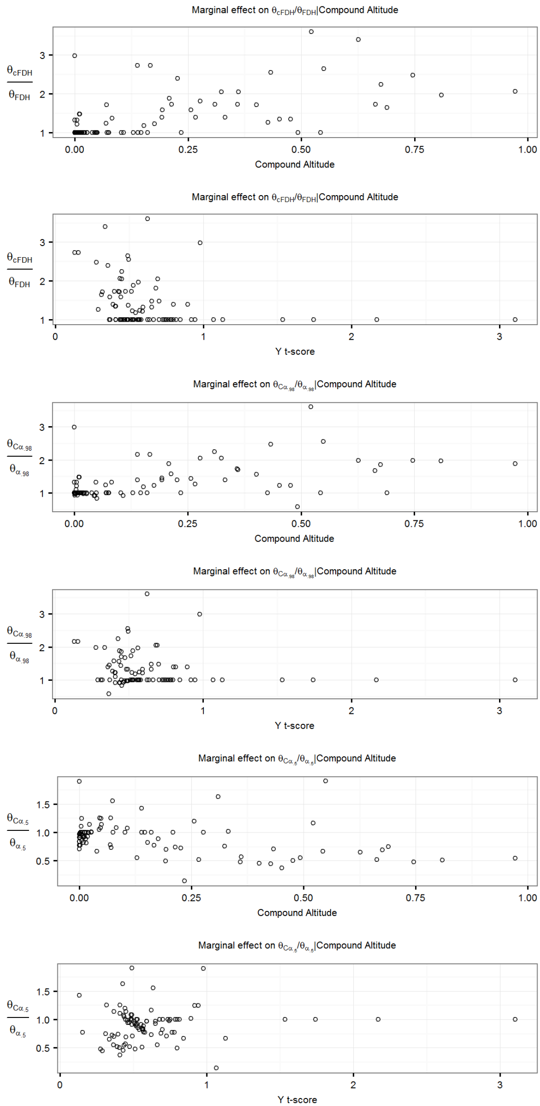


Figure 14: Crime Incidence. For each block, the first plot refers the marginal impact of z on the efficiency ratios, while the second plot shows the marginal impact of y. Blocks stand for: full frontier, robust version ($\alpha = 0.99$) and middle distribution ($\alpha = 0.50$), respectively.

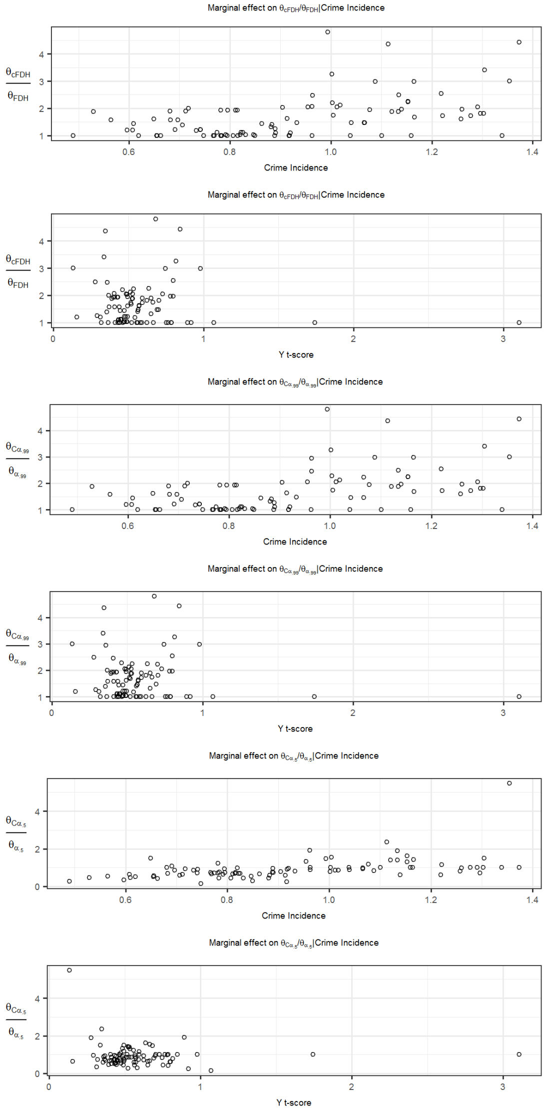
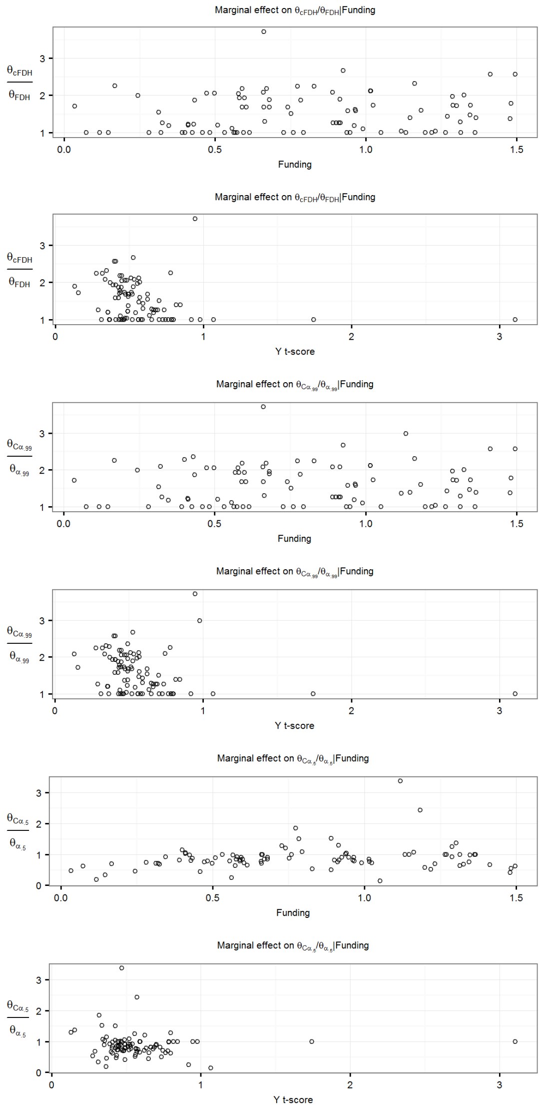


Figure 15: Funding. For each block, the first plot refers the marginal impact of z on the efficiency ratios, while the second plot shows the marginal impact of y. Blocks stand for: full frontier, robust version ($\alpha = 0.98$) and middle distribution ($\alpha = 0.50$), respectively.

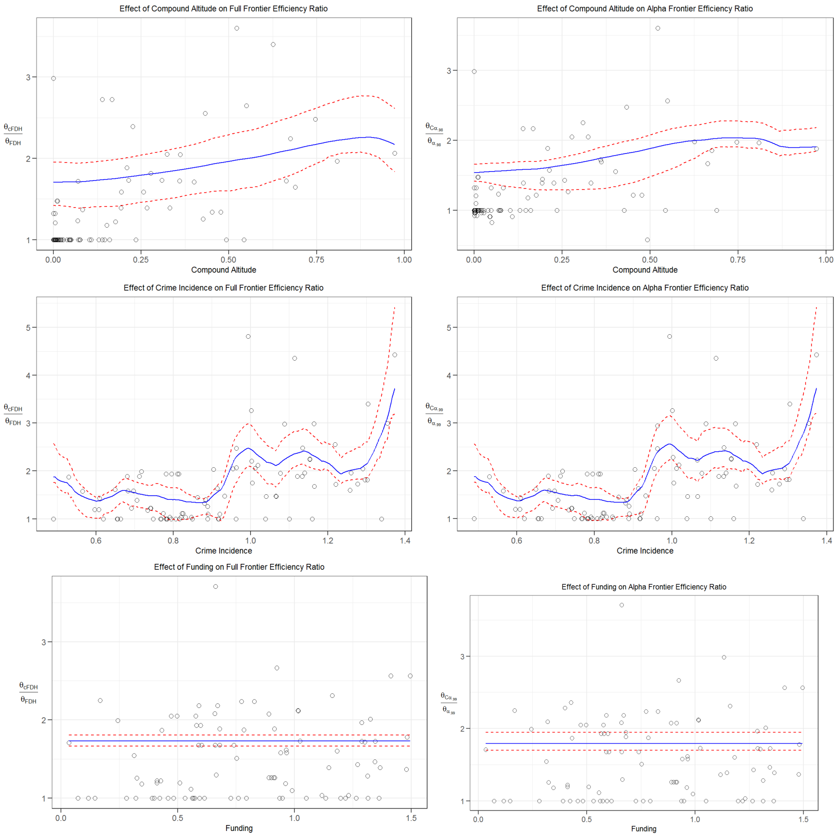


4.5.2 Confidence Intervals

Here we present and discuss the detection of the impacts of the Z external measures on conditional efficiencies. We follow the methodology introduced in Sub-Section 4.4.4. As outlined above, this analysis is meaningful since there is not interaction with the output y. Basically, we attempt to assess the significance of Z measures on the output process by applying a confidence interval built on a bootstrap procedure (Badin et al. 2012). Hence, in Figure 16 for each measure z we present its marginal effect (the continuous blue line) with the corresponding 95% confidence intervals (the dotted red lines) on both the ratio $R_I(z)$ and its robust analog $R_{\alpha,I}(z)$.

By construction, in the input-oriented scenario all the ratios are larger than one for the full frontier case. Figure 16 reveals that $Compound\ Altitude$ presents a significant effect which is slightly increasing for higher values of it. Hence, there is a modest but significant positive influence of $Compound\ Altitude$ on the attainable set of efficient levels. This means that municipalities located in areas with a high mix of altitude levels and dispersion is hampered in its provisioning activity. In addition, this effect is still present once we focus on the partial frontier. The graph for $Crime\ Incidence$ displays a higher variability, the ratio is stationary in the first half of the Z distribution and then increases towards a higher plateau. The estimates are more variable in this case but a significant and increasing effect seem to be at stake along the whole distribution; a similar pattern is shown by the robust version of the frontier. Hence, $Crime\ Incidence$ is playing an analogous role of $Compound\ Altitude$: DMUs with higher values of crime tend to be less efficient. Finally, we notice that $Funding$ does not play any role on both the full and partial frontiers. Therefore, it seems that the structure of the municipality stock of debt is not related to particular marginal effects, either as a favourable or as an unfavourable impact. Conclusions may suffer from the presence of potential outliers, however the analysis of the robust versions of the frontiers reveals that significant effects of these Z measures are basically confirmed. Clearly the presence of very high values for z may mask the true effects of this external measure on the efficiency scores. These effects seem to be reasonably disentangled by the application of the robust version of the frontier. The marginal analysis suggests the presence of two external measures, namely $Compound\ Altitude$ and $Crime\ Incidence$, which significantly impact on the output process. These measures act as unfavourable inputs, thus motivating their inclusion in the second stage analysis aimed to identify efficiency scores which are cleaned from these external effects. In addition, we observe that the stock of debt ($Funding$) presents a significant but flat impact on the efficiency scores. This means that the burden of funding constraints related to the level of debt, which incumbent municipality managers inherit from the past, does not seem to discriminate efficiency performances.

Figure 16: Bootstrapped Confidence Intervals. Plots in the rows refer to $Compound\ Altitude$, $Crime\ Incidence$ and $Funding$, respectively. Plots on the left refer to the full frontier, while those on the right stand for the robust case. Blue lines exhibit the marginal effects, while the dotted red lines stand for the 95% confidence intervals.



4.5.3 Second Stage Analysis

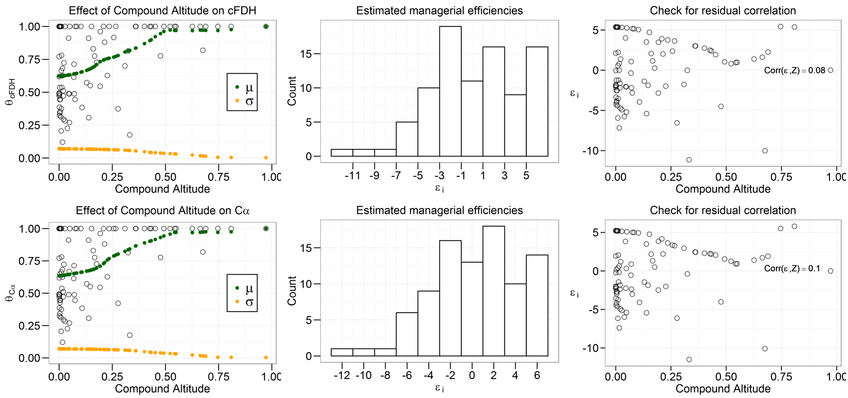
Local effects of the Z external variables on conditional measures of efficiencies that are shown in Sub-sections 4.5.1 and 4.5.2 support the second-stage analysis that we present below. We apply the nonparametric approach introduced by Badin et al. (2012) and discussed in Section 4.4.5. Basically, the conditional efficiency scores are regressed against each external variable z in a location/scale model. Using this methodology it is possible to characterize the local average effect of z on the conditional efficiency, $\mu(z)$, while also estimating locally the dispersion of its distribution, $\sigma(z)$, thus allowing for the presence of heteroscedasticity, in formula: $\theta(X, Y|Z = z) = \mu(z) + \sigma(z)\epsilon$.

Since the effects of y values on the efficient frontier is not clear and visible, we reasonably assume that partial separability is verified, thus legitimating the second stage analysis. This allows us to isolate the pure managerial efficiencies which we will express in terms of ϵ_i. Notably, unlike efficiency measures as FDH or DEA, ϵ_i of DMUs characterized by different levels of Z may be consistently ranked because the effect of the exogenous variable has been accounted for.

We provide the second stage analysis for each external measure in both its full frontier and robust version. Figures 17-19 will present these results for each z separately. In particular, we will show: i) the average (μ) effect of z and its dispersion (σ); ii) the estimated distribution of pure efficiencies (ϵ_i); iii) the correlation between pure efficiencies and the values of the selected external measure. This framework allows us to detect the impacts of z on conditional efficiency scores and to assess whether ϵ_i are appropriate measures of efficiency.

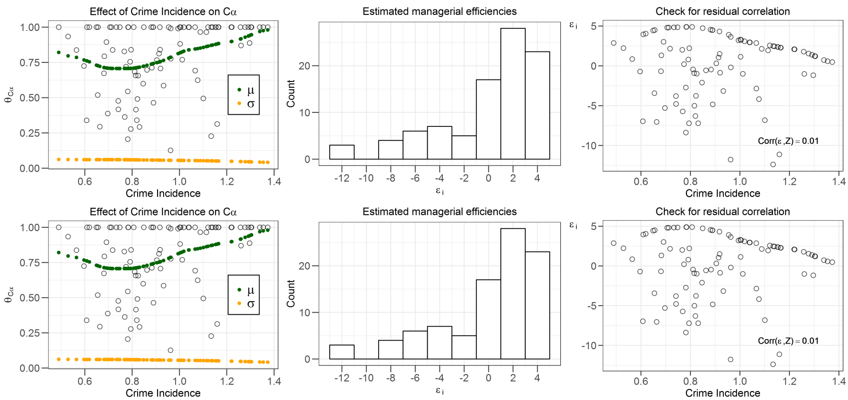
- **Compound Altitude.** Figure 17 shows that the average effect of $Compound\ Altitude$ is increasing in z, while the role of σ is minimal, being almost flat and low along the entire range of z. These behaviors seem to be persistent even in the robust version of the frontier. The histogram of the estimated pure efficiencies ϵ is shown in the middle plot of Figure 17. The distribution of ϵ_i is useful to compare municipalities once the main effects of z are eliminated. In fact, the DMUs on the left of the distribution are characterized by a lower managerial efficiency than those at the right. The ranking provided by ϵ_i indicates the emergence of a bundle of efficient municipalities, while a substantial subset of DMUs still result far from the efficient conditional frontier. It has to be noted that the comparison of the residuals is legitimate because the influence of z has been removed from the conditional scores (Sub-Section 4.5.4 will specifically consider multiple Z to gauge managerial efficiency using a multidimensional perspective). Finally, the scatterplot in the right panel of Figure 17 shows the correlation between the estimated pure efficiencies ϵ_i and the corresponding values of z_i. An appropriate second stage analysis would require this correlation to be low, meaning that ϵ_i can be used as a proxy for the pure managerial efficiency free from location and scale effects due to z. Actually, the correlation is only 0.08 in the full case (0.10 in the robust version), confirming that the regression successfully cleaned out the relationship between z_i and the conditional efficiency scores given by $Compound\ Altitude$. The Pearson correlation between these whitened scores and the original ranking based on the conditional efficiency scores is 0.87, while the Spearman rho is 0.89 (for both the full and robust versions).

Figure 17: Second Stage Analysis: Compound Altitude The first panel refers to the full frontier case, while the second one shows the effects on the robust version. Plots on the left exhibit the location and scale effect due to μ and σ; plots on the middle refer to the distribution of the managerial efficiencies ϵ_i; plots on the right show the correlation between ϵ_i and the corresponding z values.



- **Crime Incidence.** Figure 18 shows that the average effect of z is, for the full and the partial frontier alike, slightly U-shaped in the left part of the Z distribution with an increasing pattern on the right. This is coherent with the findings presented in Figure 16 of Sub-Section 4.5.2, where we observed that ratios were on a lower plateau for low values of Z and on a higher one for higher $Crime\ Incidence$ values. Moreover, the role of σ is residual, being flat and low along the entire range of z. The ranking based on the estimated ϵ_i results, therefore, similar to the one provided by the conditional efficiency scores. The Pearson correlation between whitened efficiencies and conditional efficiency scores is 0.95 for both the full and robust versions, while the Spearman rho is 0.89 and .88, respectively. Even in this case, the second-stage analysis is based on estimated pure efficiencies ϵ_i which are poorly correlated to z values (0.01 for both full and robust versions), hence, the resulting measures of managerial efficiency based on ϵ_i has been nicely whitened, thus providing a reasonable ranking of efficiency.

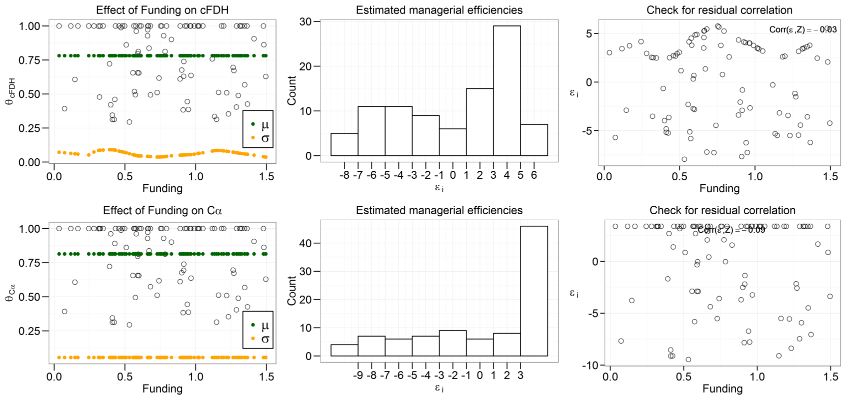
Figure 18: Second Stage Analysis: Crime Incidence The first panel refers to the full frontier case, while the second one shows the effects on the robust version. Plots on the left exhibit the location and scale effect due to μ and σ; plots on the middle refer to the distribution of the managerial efficiencies ϵ_i; plots on the right show the correlation between ϵ_i and the corresponding z values.



- **Funding.** Figure 19 exhibits the impact of $Funding$. The first plot reveals that its influence is almost negligible in terms of different values of z, being substantially flat along the entire range of it. Indeed, the bandwidth selected by means of LSCV in the second stage is large compared to the range of Z. Similar to the previous variables, the role of σ is very marginal and the picture depicted by the robust version of the frontier confirms the findings of the full frontier case. The estimated distribution of ϵ_i is, therefore, simply a standardised version of the distribution of conditional measures (Pearson correlation and Spearman rho are equal to 0.96 and 0.91 in the full frontier, while in the robust version they are almost equal to 1 and 0.95 respectively). Still, the distribution of ϵ_i is correct and the correlation between ϵ_i and the corresponding z_i values confirms the absence of a clear relationship (the correlations are -0.03 and -0.09 with respect to the full and robust case, respectively). Those units that are on the left side of the distribution of estimated ϵ_i should deserve special attention to understand which are the generating processes of their inefficiency levels. However, it seems that $Funding$ does not play the role of either favourable or unfavourable input, which seems to be an odd result if interpreted in terms of budget constraints that municipalities have to face to respect fiscal stability requirements.

The unidimensional second-stage analysis provide separate perspectives on efficiency scores cleaned from the impact of one external dimension at time. Figure 20 show rankings based on the different values of ϵ_i obtained from the second-stage analysis for both the full and partial frontiers. To increase map readability, the efficiency of the municipality is imputed to the whole province it belongs to. The visualization the managerial efficiencies belonging to different Z displays a peculiar

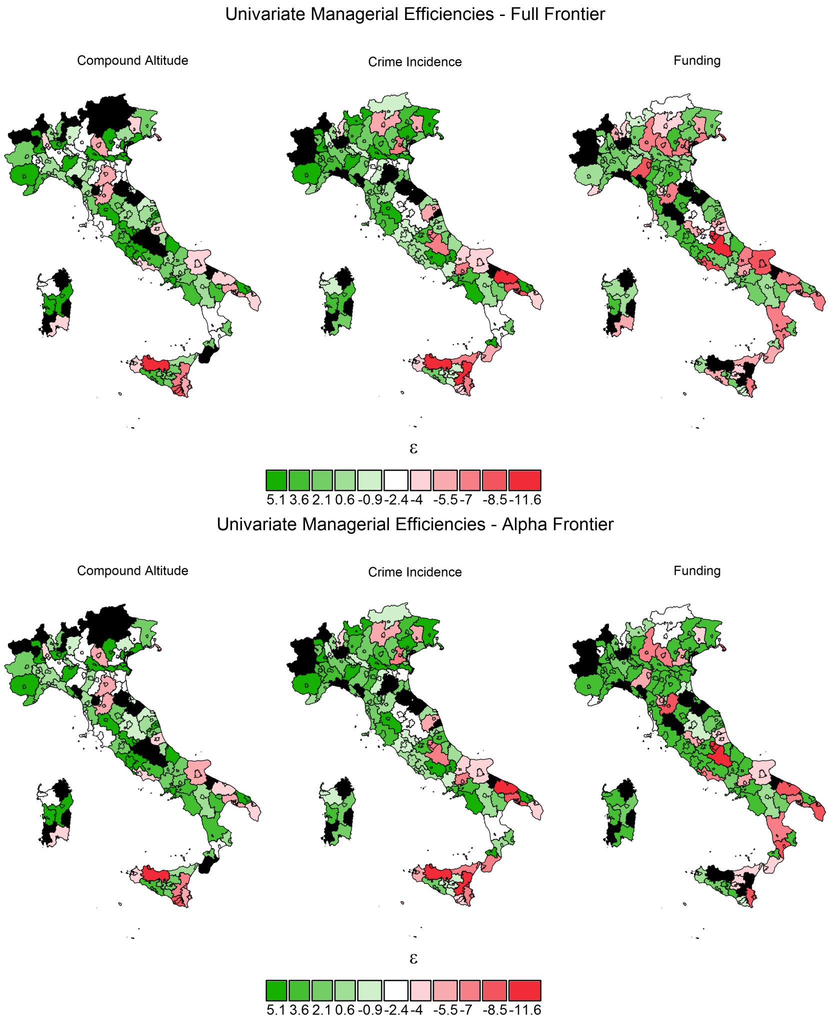
Figure 19: Second Stage Analysis: Funding The first panel refers to the full frontier case, while the second one shows the effects on the robust version. Plots on the left exhibit the location and scale effect due to μ and σ; plots on the middle refer to the distribution of the managerial efficiencies ϵ_i; plots on the right show the correlation between ϵ_i and the corresponding z values.



pattern of the efficiency distribution. While some municipalities tend to persistently rank at the top/bottom in every specification, a relevant dispersion in the core of the distribution is displayed across different exogenous conditions.

Given the relevant and diverse effects of the different exogenous variables on managerial efficiency, we present in Section 4.5.4 an attempt to account for more external variables at the same time; i.e., a multi-dimensional second-stage analysis. This scrutiny allows to identify joint effects neglected in the univariate analysis and to estimate better managerial efficiencies since exogenous effects are more thoroughly washed out. However, it has to be noted that increasing the number of environmental variables considered reduces the precision of estimates and

Figure 20: Maps of efficiency levels for single external dimension. The first panel refers to the full frontier case, while second shows the partial frontier scenario. We consider, from left to right: *Compound Altitude*, *Crime Incidence* and *Funding*. Green areas mean higher managerial efficiency, while red stands for the tail of the distribution corresponding to inefficient levels. In black are those areas discarded for the specific unidimensional analysis.



greatly increase the computational burden of the analysis.

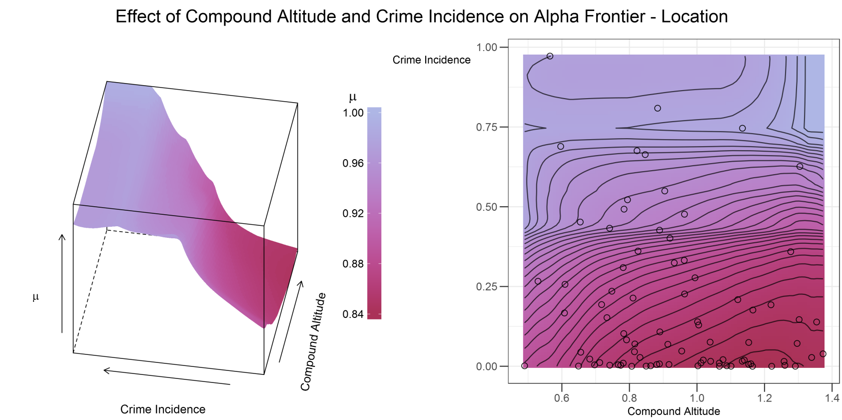
4.5.4 Bivariate Analyisis: the Joint Impact of Compound Altitude and Crime Incidence

In this Section we explore the conjoint impact of $Compound\ Altitude$ and $Crime\ Incidence$ on municipalities' efficiencies. We are not considering $Funding$ in this setting since the univariate analysis has established its modest impact. The bivariate analysis consists of a location-scale regression model, analogous to the one used in the univariate case, where both the exogenous variables are considered together. Given that the univariate analysis has detected the presence of some extremes, here we opt for the robust measure of α quantile with $\alpha = 0.98$.

Figure 21 illustrates the local effect of $Compound\ Altitude$ and $Crime\ Incidence$ on the Alpha frontier: in the upper panel the local mean is depicted while in the bottom one the local standard deviation is considered. For both panels, we report a surface graph on the left, while on the right we provide the relative contour plot. The multidimensional analysis mostly confirms the marginal ones. The local mean effect is increasing in $Crime\ Incidence$ and this behaviour seem to be affected only marginally by $Compound\ Altitude$. Firstly, for high values of $Crime\ Incidence$ the local effect is just slightly smaller for high values of $Compound\ Altitude$ than for lower ones; secondly, for high values of $Compound\ Altitude$ the rate of increase of the local average is more smoothed. Conversely, $Compound\ Altitude$ has a qualitatively different impact on the local mean across the range of $Crime\ Incidence$: for low values of the latter it affects negatively the local mean while the opposite holds for high values of $Crime\ Incidence$. However, we highlight that the distribution of sample observations is not homogeneous and the estimates of the local effect for high values $Crime\ Incidence$ and/or small values of $Compound\ Altitude$ are based on a small number of points. The scale impact is of extremely low magnitude and seems to be not influenced by $Compound\ Altitude$, while it is negatively related with $Crime\ Incidence$. Despite an increased noisiness, similar results hold for the full frontier (see Figure 27 in section 4.8).

The multidimensional specification of the location-scale model succeed in removing the effects of the two exogenous variables considered: the correlation between the managerial efficiencies and $Compound\ Altitude$ is -0.04, while the one with respect to $Crime\ Incidence$ is 0.07 (the correlation when the full frontier is considered are even lower and equal to -0.02 and 0.02, respectively). In the left panel of Figure 22 it is reported the histogram of the managerial efficiencies. The distribution is heavily skewed towards higher values, a result coherent with the richer number of exogenous factors employed in this Section (Li and Racine 2007), though entailing a reduced discriminatory power of the analysis. In the right panel we provide some summary statistics of managerial efficiencies broken down by geographical area. Given the remarkable skewness of the distribution, we will focus on the median behaviour. North West, North-East, and Center are ranked together in first position, while South and Islands follows with slightly more and slightly less than half of their efficiencies, respectively. In Figure 23 the same investigation is undertook with respect to the unconditional alpha frontier measures of efficiency. As it can be seen, not accounting for exogenous factors deeply affects the results: while the ranks of North-West and Islands remain sta-

Figure 21: Second Stage - Compound Altitude and Crime Incidence - Alpha Frontier. In the upper panel we show the location effect, while in the bottom panel we exhibit the scale effect. Plots on the left refer to the 3-d representation of the joint impact of $Compound\ Altitude$ and $Crime\ Incidence$, while on the right we show the corresponding counterplots. The bivariate analysis refers to the robust version of the frontier when $\alpha = 0.98$.



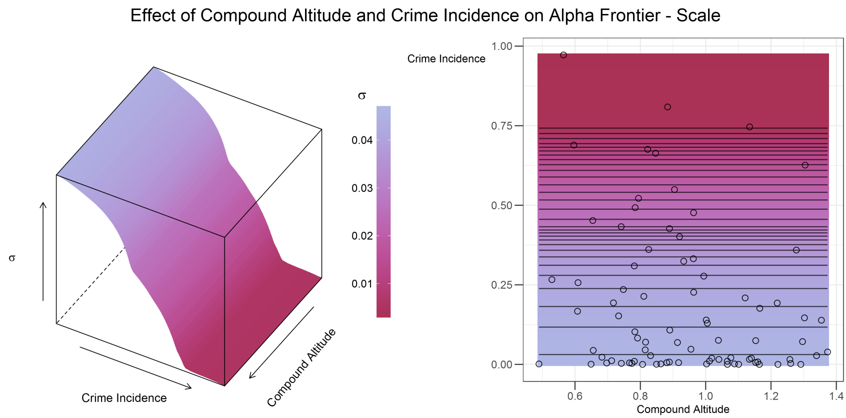
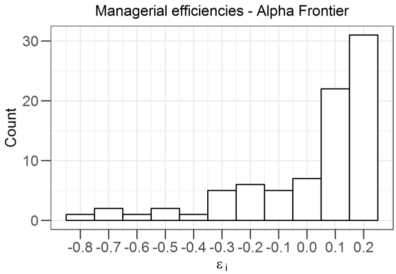
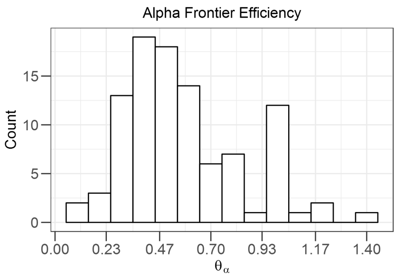


Figure 22: Distribution and Summary Statistics by Geographical Area - Partial Multidimensional.



Multidimensional Efficiency - α Frontier				
area	mean	median	sd	n.obs
North-West	0.06	0.15	0.16	16
North-East	-0.02	0.15	0.25	16
Center	0.04	0.15	0.22	17
South	0.02	0.09	0.22	21
Islands	-0.08	0.06	0.31	13

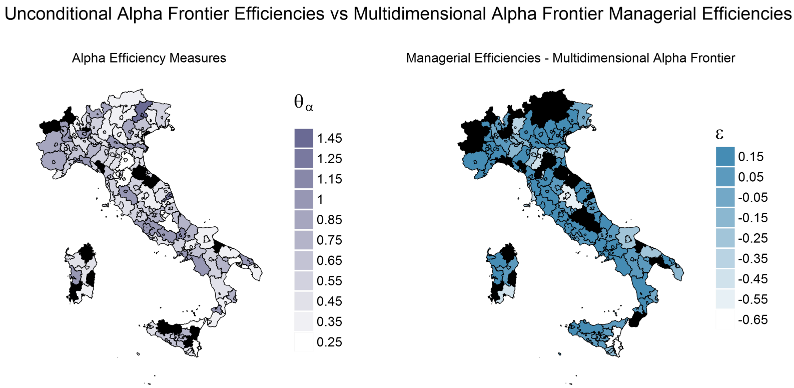
Figure 23: Distribution and Summary Statistics by Geographical Area - Partial Unconditional.



Unconditional Efficiency - α Frontier				
area	mean	median	sd	n.obs
North-West	0.66	0.61	0.23	22
North-East	0.54	0.49	0.28	21
Center	0.60	0.53	0.26	20
South	0.60	0.58	0.29	23
Islands	0.46	0.44	0.25	13

ble, the South moves in second position while the Center is now third and North-East falls in fourth position. Moreover, the histogram in the latter figure displays just a small negative skewness. To highlight the points just raised, in Figure 24 we plot on the Italian map the "pure" managerial efficiencies obtained from the multidimensional specification and the Alpha measures of efficiency (respectively, left and right panels). As it can be easily seen, the right figure has a more homogeneous distribution of efficiencies than the left one. Nevertheless, a North-South pattern is visible on the former, while is absent when unconditional Alpha frontier efficiencies is employed.

Figure 24: Alpha Efficiency Measure and "Pure" Managerial Efficiencies From Multidimensional Analysis



4.6 Discussion

Aim of Section 4.4.5 is to determine qualitatively the relation between efficiency and three exogenous variables: $Compound\ Altitude$, $Crime\ Incidence$ and $Financing$.

With respect to the first one, $Compound\ Altitude$, the implicit hypothesis is that municipalities with higher level and/or dispersion of altitude are characterized by lower efficiency scores. In the literature Boetti et al. (2012) and Da Cruz and Marques (2014) investigated the same topic using a dummy for municipality above 600m and the range of altitude respectively, but find no evidence to support a statistically significant relation with efficiency

The scientific literature provides supporting evidence of a positive relation between crime and efficiency in lo Storto (2016) (micro-crimes per capita is used as a proxy) and in Da Cruz and Marques (2014) (using crime rate in the analysis).

Finally, a rich literature studied the relationship between debt and efficiency and has provided theoretical interpretations and empirical evidence supporting both positive and negative effects. Evidence of a positive link has been rationalized in relation to capital: a larger stock of debt may be related to higher quantity and quality of the capital stock that determines efficiency gains that more than offset their cost in terms of interests. Another explanation is based on fiscal capacity considerations: municipalities with low fiscal capacity are the ones forced to rely on debt which, by reducing the stability of the financial/economic equilibrium of the entity considered, exerts a “disciplinary” effect that eventually determines an improvement of efficiency (Worthington 2000; Benito et al. 2010). Conversely, evidence of a negative effect of debt has been mainly related to the diversion of resources from productive means to unproductive interest and amortization costs it entails.(Geys 2006; Geys and Moesen 2009; Ashworth et al. 2014; Da Cruz and Marques 2014; Cordero and Pedraja-chaparro 2016). Finally, numerous studies provided no evidence to support a relation between debt and efficiency (Balaguer-Coll et al. 2007; Revelli and Tovmo 2007; Benito et al. 2007; Balaguer-Coll and Prior 2009).

As anticipated in section 4.2, the most closely related study to the present one is lo Storto (2016). Contrary to our findings, a positive relationship between crimes and efficiency is presented in the article. The

Author interprets this result in the light of the level of development of a municipality: a higher level of urban micro-criminality and a reduced capability of the territory to produce economic value added is related to lower spending that leads to improved efficiency. Given that the latter study also reports the whole set of efficiency scores, it is possible to directly investigate its relations with all the measures here developed: unconditional, conditional-univariates and conditional-multivariate. In particular, we undertook an analysis of the correlations of both constant return to scale (crs) and variable return to scale (vrs) DEA measures presented in lo Storto (2016) with the ones here developed. From this investigation it can be concluded that the measures of the present study lack a significant correlation to the ones developed by lo Storto (2016) with the exception of conditional-univariate scores considering $Crime\ Incidence$, whose Pearson and Spearman coefficients are the more relevant (maximal with respect to the vrs-DEA) but negative and equal to -0.209 in both cases (p-values being 0.05 and 0.04 respectively). In fact, even focusing on the correlation coefficient of higher magnitude (the Spearman one) and considering the more "related" efficiency measure (crs or vrs), correlation is low and insignificant: unconditionl/vrs-DEA -0.117 with a p-value of 0.25, conditional-univariate-$Financing$/crs-DEA -0.084 with a p-value of 0.42, conditional-multivariate/vrs-DEA -0.06 with a p-value of 0.57 and conditional-univariate-$Compound\ Altitude$/vrs-DEA -0.05 with a p-value of 0.66. Despite the similarities with our work, we argue that our contribution provides an improvement in the completeness and reliability of the efficiency evaluation of major Italian municipalities. From a methodological point of view, the present analysis has two main strengths. First, outliers treatment has been thoroughly performed using an improved version of the methodology proposed by Simar (1996) and Daraio and Simar (2007a) that is able to account for the conditional nature of the efficiency estimates (see the discussion Section 4.4.7 for the relevance of this point). Second, the methodology adopted relies only on the less-restrictive partial-separability condition to hold, a requirement whose fulfilment is ensured in Section 4.5.1. Moreover, we argue that the composite indicator of municipality production activity here developed provides a more comprehensive and realistic measure of outputs relative to the narrower set of variables used in previous studies.

In Boetti et al. (2012), an efficiency assessment of Piemonte's municipalities has been performed. Unfortunately, in this case it is not possible to directly compare efficiencies, however, the article provides an analysis of the effect of altitude on efficiency by including a dummy variable for high-altitude observations. Altough the article presents no evidence of a relation between altitude and efficiency, we are able to identify a negative link between the two in our chapter. While the remarks on methodology and data apply in this case too, given the different set on municipalities under analysis the two result are not mutually exclusive.

4.7 Conclusion

This study applies conditional frontier models to assess the efficiency of major Italian municipalities for the year 2011. Leveraging on a novel data set with an unprecedented coverage of goods and services offered by municipalities, we introduce a composite indicator of local government output that provides a comprehensive evaluation of the wide range of activities performed by municipalities. Using accounting data, we compute a cost-related input measure by matching the functions considered in the output indicator with the corresponding expenses. Finally, the composite input-output indicator is used to evaluate managerial performances conditional to the exogenous conditions faced by municipalities.

The choice of the external/exogenous measures is based on intuitive motivations. Obviously, it seems not fair to compare municipalities characterized by different physical conditions and an adjusted measure of altitude (i.e. the $Compound\ Altitude$) is, therefore, included to account for topographic diversity. Similarly, the social context may influence the feasible policies implementable by the local manager plausibly impacting the performances of the municipality. For this reason, we include the level of $Crime\ Incidence$ to approximate the social environmental conditions. Finally, we introduce the historical level of the stock of debt (i.e. $Funding$) to verify whether this constraint plays a role in determining municipality efficiency.

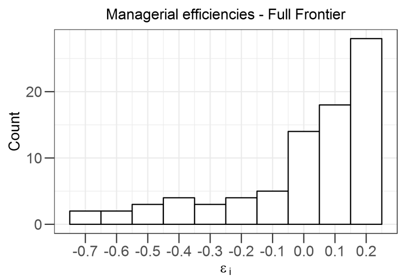
Our findings show that $Compound\ Altitude$ acts as an unfavourable input: municipalities characterized by a higher mix of altitude level and dispersion are less likely to reach the efficient frontier *ceteris paribus* the combination of inputs and outputs. A slightly negative impact is also observed for $Crime\ Incidence$, a result implying that additional efforts are needed by municipalities with high values of this z to achieve efficient performances. Conversely, we find no evidence of an effect of $Funding$ on municipality efficiency.

The finding of a negative impact of $Compound\ Altitude$ on efficiency is novel in the literature and likely obtained thanks to the improvements introduced by the present chapter, in terms of both data and methodology, with respect to previous studies. Similarly, evidence of a negative relationship between crime and performance has never been presented before and we argue it could serve as a starting point for a fruitful debate on the topic. Finally, our investigation on the role of $Financing$, while not entirely novel, it is the first one referring to the Italian case.

The methodology adopted in the analysis, i.e. the conditional frontier model, has remarkable strengths and weaknesses. From one hand, the results obtained are based on an extremely narrow set of assumptions and can therefore be considered remarkably robust. From the other hand, the high computational burden needed to perform the analysis limits the dimensionality of the exogenous variables that can be considered. Taking into account those traits is vital to a correct interpretation of the presented results. "Pure" managerial efficiencies here provided have been washed out from the effects of just two of the plausibly many exogenous factors affecting it. Hence, while representing a better measure relative to the unconditional one (and to the univariate ones too), the "pure" managerial efficiency provided in the bivariate exercise should be taken with caution if exploited for policy guidance. Still, we remark that the inclusion of just two exogenous factors has major effects on the estimated managerial efficiencies and, therefore, an analysis failing to account for this aspect would be fundamentally flawed.

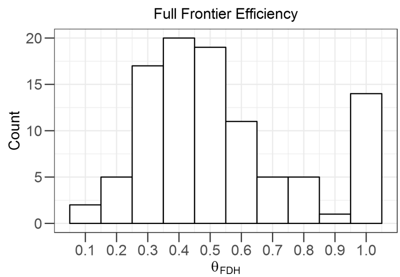
Theoretical and implementation improvements that would allow to include a wider set of environmental variables in the analysis is of primary interest for further research. Indeed, the analytical framework adopted by the chapter entails computational complexities that limit the set of exogenous variables that could be included in the analysis while methodological constraints call for a strict interpretation of the "exogeneity condition" in the selection of environmental variables. Furthermore, an additional avenue for future research lies in the expansion of the analysis along the time dimension in order to investigate the dynamics of efficiency also in relation to governmental and local policies. Finally, an extension of the current work to the whole set of Italian municipalities would both improve estimation reliability and provide useful information to the policymaker.

4.8 Supporting Information



Multidimensional Efficiency - Full Frontier				
area	mean	median	sd	n.obs
North-West	0.05	0.09	0.19	16
North-East	-0.05	0.05	0.32	16
Center	0.04	0.10	0.26	17
South	0.02	0.09	0.22	21
Islands	-0.07	0.03	0.31	13

Figure 25: Distribution and Summary Statistics by Geographical Area - Full Multidimensional



Unconditional Efficiency - Full Frontier				
area	mean	median	sd	n.obs
North-West	0.62	0.55	0.23	22
North-East	0.50	0.47	0.23	21
Center	0.57	0.50	0.25	20
South	0.53	0.48	0.27	23
Islands	0.41	0.38	0.23	13

Figure 26: Distribution and Summary Statistics by Geographical Area - Full Unconditional

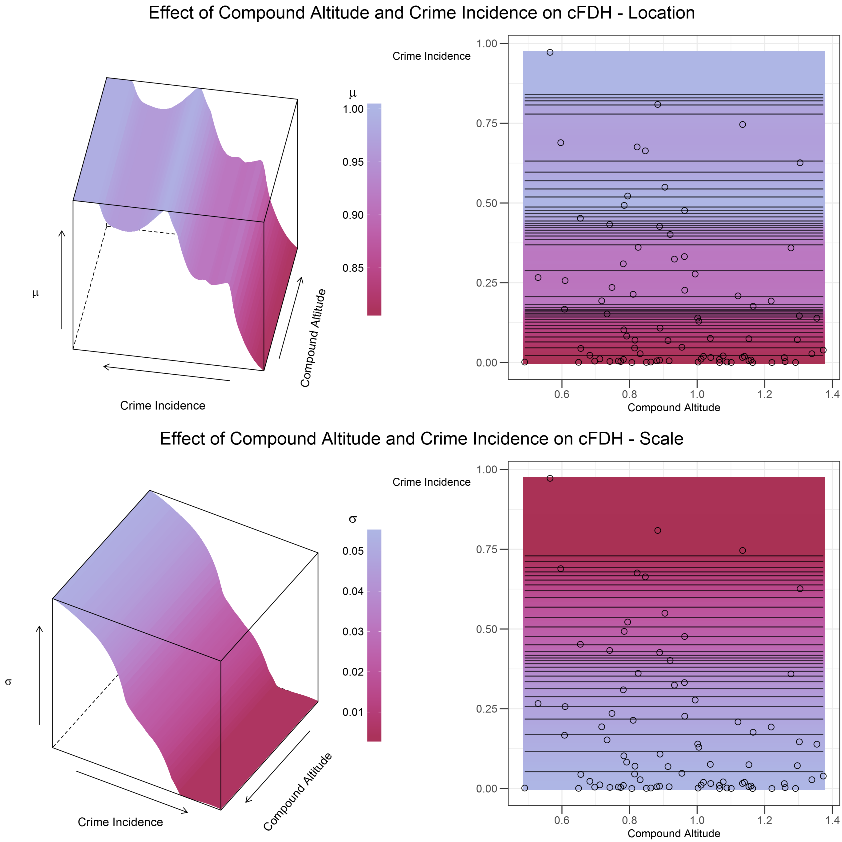


Figure 27: Second Stage - Compound Altitude and Crime Incidence - Full Frontier

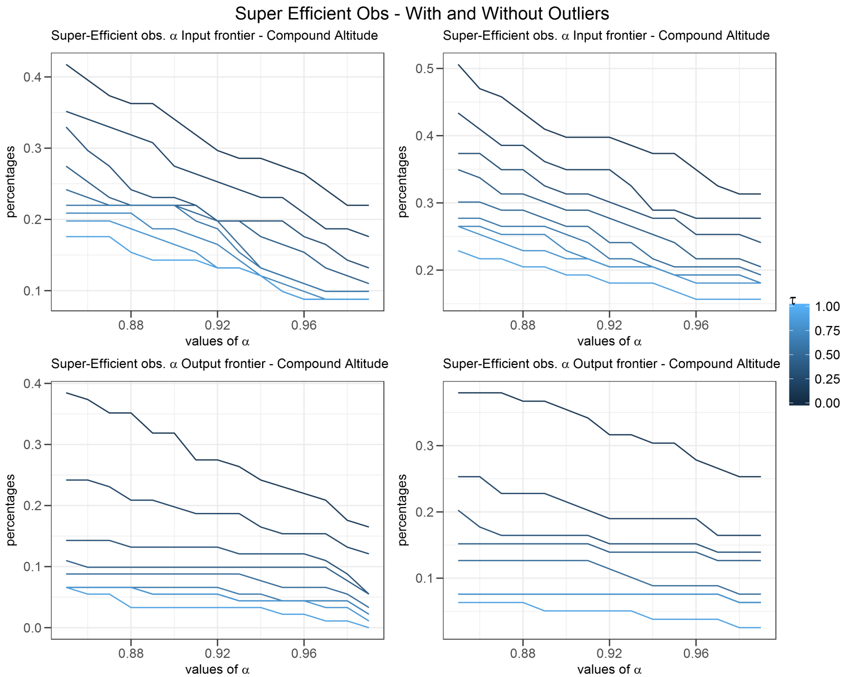


Figure 28: Outlier Detection and choice of optimal α - Compound Altitude (Left panel whole sample, Right panel without outliers)

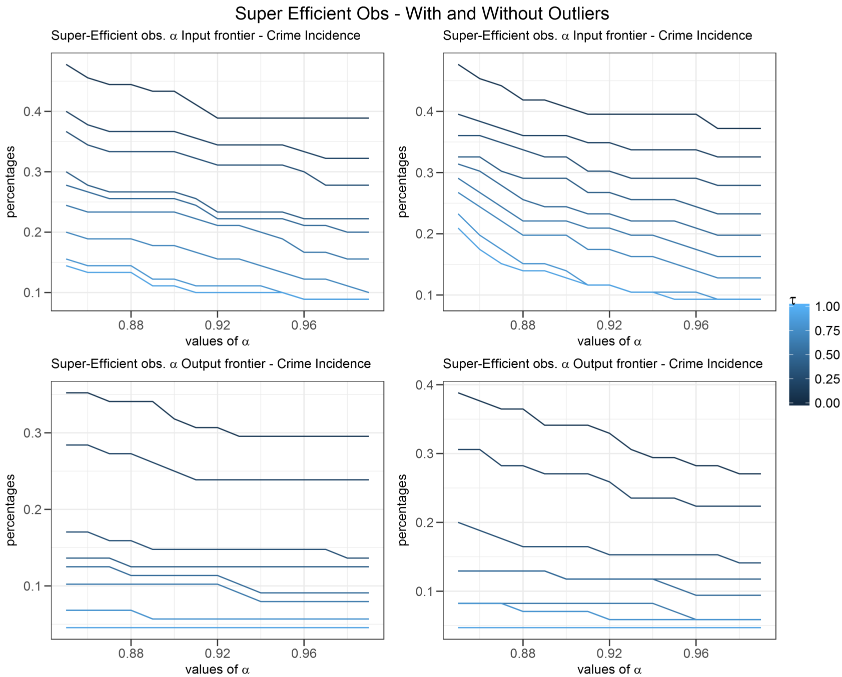


Figure 29: Outlier Detection and choice of optimal α - Crime Incidence (Left panel whole sample, Right panel without outliers)

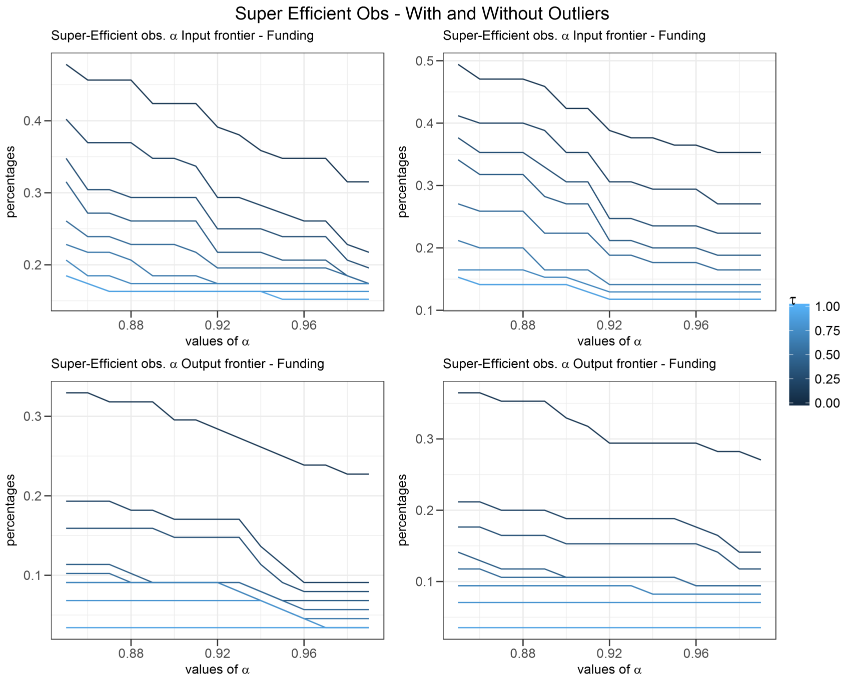


Figure 30: Outlier Detection and choice of optimal α - Funding (Left panel whole sample, Right panel without outliers)

References

Afonso, A., Fernandes, S. 2006. Measuring local government spending efficiency: evidence for the lisbon region. Regional Studies, 40, 39–53. 80, 102, 103

Afonso, A., Fernandes, S. 2008. Assessing and explaining the relative efficiency of local government. The Journal of Socio-Economics, 37, 1946–1979. 80, 102, 103

Akerlof, G. A. 1982. Labor contracts as partial gift exchange. The Quarterly Journal of Economics, 543–569. 3

Al-Nowaihi, A., Pyle, D. 2000. Income tax evasion: a theoretical analysis. In A. MacDonald, D. Pyle eds. Illicit Activity: The Economics of Crime and Tax Fraud, Ashgate, Aldershot, 249–266. 43, 61

Allingham, M. G., Sandmo, A. 1972. Income tax evasion: A theoretical analysis. Journal of Public Economics, 1, 323—338. 1, 6, 35

Alm, J. 1988a. Compliance costs and the tax avoidance-tax evasion decision. Public finance quarterly : PFQ, 16, 31–66. 9, 25

Alm, J. 1988b. Compliance costs and the tax avoidance tax evasion decision. Public Finance Quarterly, 16, 31–66. 40

Alm, J., Bahl, R., Murray, M. N. 1990. Tax structure and tax compliance.. The Review of Economics and Statistics, 72, 603–13. 18, 23, 40

Alm, J., Finlay, K. 2013. Who benefits from tax evasion? Economic analysis and policy, 43, 139–154. 31

Alm, J., Jackson, B. R., McKee, M. 1993. Fiscal exchange, collective decision institutions, and tax compliance. Journal of Economic Behavior & Organization, 22, 285–303. 4

Alm, J., McCallin, N. J. 1990. Tax avoidance and tax evasion as a joint portfolio choice. Public Finance= Finances Publiques, 45, 193–200. 9, 24, 40

Alm, J., McClelland, G. H., Schulze, W. D. 1992. Why do people pay taxes? Journal of public Economics, 48, 21–38. 3, 4, 27

Alm, J., Sanchez, I., De Juan, A. 1995. Economic and noneconomic factors in tax compliance. Kyklos, 48, 1–18. 3

Alm, J., Torgler, B. 2006. Culture differences and tax morale in the united states and in europe. Journal of economic psychology, 27, 224–246. 3

Ammons, D. 2014. Municipal Benchmarks: Assessing Local Perfomance and Establishing Community Standards. Routledge. 88

Ashworth, J., Geys, B., Heyndels, B., Wille, F. 2014. Competition in the political arena and local government performance. Applied Economics, 46, 2264–2276. 80, 127

Athanassopoulos, A. D., Triantis, K. P. 1998. Assessing aggregate cost efficiency and the related policy implications for greek local municipalities. Infor, 36, p. 66. 80

Bădin, L., Daraio, C. 2011. Explaining efficiency in nonparametric frontier models: recent developments in statistical inference. In Exploring research frontiers in contemporary statistics and econometrics, Springer, 151–175. 98

Badin, L., Daraio, C., Simar, L. 2010. Optimal bandwidth selection for conditional efficiency measures: a data-driven approach. European Journal of Operational Research, 201, 633–640. 94, 95, 96, 101, 102

Badin, L., Daraio, C., Simar, L. 2012. How to measure the impact of environmental factors in a nonparametric production model. European Journal of Operational Research, 223, 818–833. 79, 106, 112, 115

Bădin, L., Daraio, C., Simar, L. 2014. Explaining inefficiency in nonparametric production models: the state of the art. Annals of Operations Research, 214, 5–30. 81, 97, 98, 100

Balaguer-Coll, M. T., Prior, D. 2009. Short- and long-term evaluation of efficiency and quality. An application to Spanish municipalities. Applied Economics, 41, 2991–3002. 78, 80, 127

Balaguer-Coll, M. T., Prior, D., Tortosa-Ausina, E. 2007. On the determinants of local government performance: A two-stage nonparametric approach. European Economic Review, 51, 425–451. 80, 88, 127

Balaguer-Coll, M. T., Prior, D., Tortosa-Ausina, E. 2010. Decentralization and efficiency of local government. The Annals of regional science, 45, 571–601. 80

Balassone, F., Zotteri, S. 2001. Il Patto di stabilità interno due anni dopo: norme più morbide, risultati più deludenti. Economia Pubblica, 6. 82

Baldry, J. C. 1986. Tax evasion is not a gamble. Economics Letters, 22, 333–335. 43

Barberis, N., Huang, M., Thaler, R. H. 2006. Individual preferences, monetary gambles, and stock market participation: A case for narrow framing. American Economic Review, 96, 1069–1090. 11

Barker, W. B. 2009. The ideology of tax avoidance. Loyola University Chicago Law Journal, 40, 229–251. 12

Baron, D. P. 1994. Electoral Competition with Informed and Uniformed Voters. The American Political Science Review, 88, 33–47. 77

Barone, G., Mocetti, S. 2011. Tax morale and public spending inefficiency. International Tax and Public Finance, 18, 724–749. 4

Battalio, R. C., Kagel, J. H., Jiranyakul, K. 1990. Testing between alternative models of choice under uncertainty: Some initial results. Journal of risk and uncertainty, 3, 25–50. 11

Becker, G. S. 1968. Crime and punishment: An economic approach. Journal for Political Economy, 76, 169–2–17. 35

Becker, W., Heinz-Jürgen, B., Sleeking, S. 1987. The impact of public transfer expenditures on tax evasion. Journal of Public Economics, 34, 243–252. 4

Beetsma, R., Uhlig, H. 1999. An analysis of the stability and growth pact. Economic Journal, 109, 546–571. 82

Benabou, R., Tirole, J. 2003. Intrinsic and extrinsic motivation. The Review of Economic Studies, 70, 489–520. 3

Benartzi, S., Thaler, R. H. 1995. Myopic loss aversion and the equity premium puzzle. The Quarterly Journal of Economics, 110, 73–92. 11

Benito, B., Bastida, F., García, J. A. 2010. Explaining differences in efficiency: an application to spanish municipalities. Applied economics, 42, 515–528. 80, 127

Benito, B., Bastida, F., García, J. A. 2007. The determinants of efficiency in municipal governments. Applied Economics, 4, 1–14. 80, 127

Benjamini, Y., Maital, S. 1985. Optimal tax evasion and optimal tax evasion policy: behavioural aspects. In W. Gaertner, A. Wenig eds. The Economics of the Shadow Economy, benjamini1985, Springer-Verlag, Chap. Optimal ta, 245–264. 43

Bernasconi, M., Corazzini, L., Seri, R. 2014. Reference dependent preferences, hedonic adaptation and tax evasion: Does the tax burden matter?. Journal of Economic Psychology, 40, 103–118. 61

Besley, T., Case, A. 1995. Incumbent Behavior: Vote-Seeking, Tax-Setting, and Yardstick Competition. The American Economic Review, 85, 25–45. 77

Besley, T., Coate, S. 2003. Centralized versus decentralized provision of local public goods: A political economy approach. Journal of Public Economics, 87, 2611–2637. 77

Besley, T., Smart, M. 2007. Fiscal restraints and voter welfare. Journal of Public Economics, 91, 755–773. 77

Blackorby, C., Lady, G., Nissen, D., Russell, R. R. 1970. Homothetic separability and consumer budgeting. Econometrica: Journal of the Econometric Society, 468–472. 8

Boame, A. K. 2004. The technical efficiency of canadian urban transit systems. Transportation Research Part E: Logistics and Transportation Review, 40, 401–416. 80

Boetti, L., Piacenza, M., Turati, G. 2012. Decentralization and local governments' performance: how does fiscal autonomy affect spending efficiency? FinanzArchiv: Public Finance Analysis, 68, 269–302. 80, 126, 129

Bordignon. 2001. Optimal Regional Redistribution under Asymmetric Information. American Economic Review, 91, 709–723. 76

Bosch, N., Pedraja, F., Suárez-Pandiello, J. 2000. Measuring the efficiency of spanish municipal refuse collection services. Local Government Studies, 26, 71–90. 79

Bouckaert, G. 1992. Productivity analysis in the public sector: The case of the fire service. International Review of Administrative Sciences, 58, 175–200. 79

Brennan, G., Buchanan, J. M. 1980. The power to tax: Analytic foundations of a fiscal constitution. Cambridge University Press. 77

Brugnano, C., Rapallini, C. 2009. Il Patto di Stabilità Interno per i Comuni: una valutazione con i certificati dei conti consuntivi. Economia Pubblica, 1-2, 57–89. 82

Buchanan, J. M. 1967. Public Finance in Democratic Process. Fiscal Institutions and Individual Choice. Chapel Hill, North Carolina, North Carolina University Press. 4

Buchanan, J. M. 1976. Taxation in fiscal exchange. Journal of Public Economics, 6, 31–35. 4

Buettner, T., Ruf, M. 2007. Tax incentives and the location of FDI: Evidence from a panel of German multinationals. International Tax and Public Finance, 14, 151–164. 77

Buti, M., Eijffinger, S., Franco, D. 2003. Revisiting EMU's Stability Pact: A Pragmatic Way Forward. Oxford Review of Economic Policy, 19, 100–111. 82

Byrnes, J., Dollery, B. 2002. Do economies of scale exist in australian local government? a review of the research evidence1. Urban Policy and Research, 20, 391–414. 88

Camerer, C., Babcock, L., Loewenstein, G., Thaler, R. 1997. Labor supply of new york city cabdrivers: One day at a time. The Quarterly Journal of Economics, 112, 407–441. 11

Carrington, R., Puthucheary, N., Rose, D., Yaisawarng, S. 1997. Performance measurement in government service provision: the case of police services in new south wales. Journal of Productivity Analysis, 8, 415–430. 79

Cazals, C., Florens, J.-P., Simar, L. 2002. Nonparametric frontier estimation: a robust approach. Journal of econometrics, 106, 1–25. 92, 104

Chander, P. 2007. Income tax evasion and the fear of ruin. Economica, 74, 315–328. 37, 46

Chander, P., Wilde, L. L. 1998. A general characterization of optimal income tax enforcement. Review of Economic Studies, 65, 165–183. 36, 39, 45, 63, 68

Charnes, A., Cooper, W. W., Rhodes, E. 1979. Measuring the efficiency of decision-making units. European Journal of Operational Research, 3, p. 339. 77

Chetty, R. 2009. Is the taxable income elasticity sufficient to calculate deadweight loss? the implications of evasion and avoidance. American Economic Journal: Economic Policy, 1, 31–52. 43

Chiades, P., Mengotto, V. 2015. Il calo degli investimenti nei Comuni tra Patto di stabilità interno e carenza di risorse. Economia Pubblica, 2, 5–44. 82

Chiappori, P.-A., Paiella, M. 2011. Relative risk aversion is constant: Evidence from panel data. Journal of the European Economic Association, 9, 1021–1052. 13

Chiarini, B., Marzano, E., Schneider, F. 2013. Tax rates and tax evasion: An empirical analysis of the long-run aspects in Italy. European Journal of Law and Economics, 35, 273–293. 68

Christiansen, V. 1980. Two comments on tax evasion. Journal of Public Economics, 13, 389–393. 6, 8, 25, 35

Citrin, J. 1979. Do people want something for nothing: Public opinion on taxes and government spending. National Tax Journal, 32, 113–129. 69

Cobham, A. 2005. Tax evasion, tax avoidance and development finance. Queen Elizabeth House, University of Oxford, 129. 35

Committee of Public Accounts. 2013. Tax Avoidance: Tackling Marketed Avoidance Schemes. London, The Stationery Office Limited. 8, 14, 41

Cordero, M., Pedraja-chaparro, F. 2016. Efficiency assessment of Portuguese municipalities using a conditional nonparametric approach. Mimeo. 80, 127

Cremer, H., Gahvari, F. 1995. Tax evasion and the optimum general income tax. Journal of Public Economics, 60, 235–249. 45

Cummings, R. G., Martinez-Vazquez, J., McKee, M., Torgler, B. 2009. Tax morale affects tax compliance: Evidence from surveys and an artefactual field experiment. Journal of Economic Behavior & Organization, 70, 447–457. 3, 4

Da Cruz, N. F., Marques, R. C. 2014. Revisiting the determinants of local government performance. Omega, 44, 91–103. 80, 126, 127

Damjanovic, T., Ulph, D. 2010. Tax progressivity, income distribution and tax non-compliance. European Economic Review, 54, 594–607. 14, 41

Daouia, A., Gijbels, I. 2011a. Estimating frontier cost models using extremiles. In Exploring Research Frontiers in Contemporary Statistics and Econometrics, Springer, 65–81. 96

Daouia, A., Gijbels, I. 2011b. Robustness and inference in nonparametric partial frontier modeling. Journal of Econometrics, 161, 147–165. 96

Daouia, A., Simar, L. 2007. Nonparametric efficiency analysis: a multivariate conditional quantile approach. Journal of Econometrics, 140, 375–400. 95

Daraio, C., Simar, L. 2005. Introducing environmental variables in nonparametric frontier models: a probabilistic approach. Journal of productivity analysis, 24, 93–121. 2, 79, 91, 94

Daraio, C., Simar, L. 2006. A robust nonparametric approach to evaluate and explain the performance of mutual funds. European Journal of Operational Research, 175, 516–542. 96

Daraio, C., Simar, L. 2007a. Advanced Robust and nonparametric methods in efficiency analysis: methodology and applications. Springer Science & Business Media. 91, 95, 96, 100, 102, 104, 129

Daraio, C., Simar, L. 2007b. Conditional nonparametric frontier models for convex and nonconvex technologies: a unifying approach. Journal of Productivity Analysis, 28, 13–32. 2, 78, 79, 92

Daraio, C., Simar, L., Wilson, P. W. 2010. Testing whether two-stage estimation is meaningful in non-parametric models of production. Institut de Statistique, UCL, Belgium. 79, 97

De Borger, B., Kerstens, K. 1996. Cost efficiency of belgian local governments: A comparative analysis of fdh, dea, and econometric approaches. Regional Science and Urban Economics, 26, 145–170. 80, 85

De Sousa, M. D. C. S., Stošić, B. 2005. Technical efficiency of the brazilian municipalities: correcting nonparametric frontier measurements for outliers. Journal of Productivity analysis, 24, 157–181. 80

De Witte, K., 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, 319–327. 80

Debreu, G. 1951. The coefficient of resource utilization. Econometrica: Journal of the Econometric Society, 273–292. 92

Dell'Anno, R. 2009. Tax evasion, tax morale and policy maker's effectiveness. The Journal of Socio-Economics, 38, 988–997. 43

Deprins, D., Simar, L., Tulkens, H. 1984. Measuring labor inefficiency in post offices. In P. P. Tulkens, H. eds. The Performance of Public Enterprises: Concepts and measurements, Amsterdam, North-Holland, 243–267. 78

Dewatripont, M., Jewitt, I., Tirole, J. 1999. The economics of career concerns, part i: Comparing information structures. The Review of Economic Studies, 66, 183–198. 77

Dhami, S., Al-Nowaihi, A. 2007. Why do people pay taxes? Prospect theory versus expected utility theory. Journal of Economic Behavior & Organization, 64, 171–192. 43

Diez-Ticio, A., Mancebon, M.-J. 2002. The efficiency of the spanish police service: an application of the multiactivity dea model. Applied Economics, 34, 351–362. 79

Dorfman, R. P., Samuelson, A., Solow, R. M. 1958. Linear Programming and Economic Analysis. McGraw Hill Text. 77

Doumpos, M., Cohen, S. 2014. Applying data envelopment analysis on accounting data to assess and optimize the efficiency of greek local governments. Omega, 46, 74–85. 80, 88, 103

EC Commission. 1990. One market, one money. European Economy, 44. 82

EC Commission. 1996. Stability Pact for Ensuring Budgetary Discipline in Stage Three of EMU : Explanatory memorandum to the Commission Proposal for a Council Regulation on the strengthening of the surveillance and coordination of budgetary positions and on speeding up and cla.Technical report, Document COM(96). 82

EC Commission: The Commission Services. 1995. Technical Note:Method for the cyclical adjustment of budget balances. European Economy, 60, 35–90. 82

Emrouznejad, A., Parker, B. R., Tavares, G. 2008. Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in dea. Socio-economic planning sciences, 42, 151–157. 79

Erard, B., Feinstein, J. S. 1994. Honesty and Evasion in the tax compliance game. The Rand journal of economics, 25, 1–19. 42

Fack, G., Landais, C. 2010. Are tax incentives for charitable giving efficient? evidence from france. American Economic Journal: Economic Policy, 2, 117–141. 40

Fan, J., Gijbels, I. 1996. Local polynomial modelling and its applications: monographs on statistics and applied probability 66. 66, CRC Press. 98, 101

Färe, R., Grosskopf, S., Lovell, C. K. 2013. The measurement of efficiency of production. 6, Springer Science & Business Media. 92

Farrell, M. J. 1957. The measurement of productive efficiency. Journal of the Royal Statistical Society. Series A (General), 120, 253–290. 77, 92

Feinstein, J. S. 1991. An econometric analysis of income tax evasion and its detection. The RAND Journal of Economics, 22, 14–35. 68

Feld, L. P., Frey, B. S. 2007. Tax compliance as the result of a psychological tax contract: The role of incentives and responsive regulation. Law & Policy, 29, 102–120. 3

Feldstein, M. 1999. Tax avoidance and the deadweight loss of the income tax. Review of Economics and Statistics, 81, 674–680. 40

Feldstein, M. 2005. The euro and the stability pact. Journal of Policy Modeling, 27, 421–426. 82

Ferejohn, J. 1986. Incumbent performance and electoral control. Public Choice, 50, 5–25. 77

Flores, R. G., Aguilera, R. V. 2007. Globalization and location choice: an analysis of US multinational firms in 1980 and 2000. Journal of International Business Studies, 38, 1187–1210. 78

Frey, B. S., Oberholzer-Gee, F. 1997. The cost of price incentives: An empirical analysis of motivation crowding-out. The American economic review, 87, 746–755. 3

Frey, B. S., Torgler, B. 2007. Tax morale and conditional cooperation. Journal of Comparative Economics, 35, 136–159. 4

García-Sánchez, I. 2008. The performance of spanish solid waste collection. Waste Management & Research, 26, 327–336. 79

Garcia-Sanchez, I.-M. 2009. Measuring the efficiency of local police force. European Journal of Law and Economics, 27, 59–77. 80

Geys, B. 2006. Looking across borders: A test of spatial policy interdependence using local government efficiency ratings. Journal of Urban Economics, 60, 443–462. 80, 127

Geys, B., Heinemann, F., Kalb, A. 2010. Voter involvement, fiscal autonomy and public sector efficiency: evidence from german municipalities. European Journal of Political Economy, 26, 265–278. 80

Geys, B., Moesen, W. 2009. Exploring sources of local government technical inefficiency: Evidence from flemish municipalities. Public Finance and Management, 9, p. 1. 127

Gideon, M. 2015. Do individuals perceive income tax rates correctly? Public Finance Review. 8

Gimnez, V. M., Prior, D. 2007. Long-and short-term cost efficiency frontier evaluation: Evidence from spanish local governments. Fiscal Studies, 28, 121–139. 80

Gneezy, U., Potters, J. 1997. An Experiment on Risk Taking and Evaluation Periods. The Quarterly Journal of Economics, 112, 631–645. 11

Gordon, J. P. 1989. Individual morality and reputation costs as deterrents to tax evasion. European economic review, 33, 797–805. 7, 43

Grossman, P. J., Mavros, P., Wassmer, R. W. 1999. Public sector technical inefficiency in large us cities. Journal of Urban Economics, 46, 278–299. 80

Gruber, J., Saez, E. 2002. The elasticity of taxable income: evidence and implications. Journal of public Economics, 84, 1–32. 40

Gupta, S., Kumar, S., Sarangi, G. K. 2012. Measuring the performance of water service providers in urban india: implications for managing water utilities. Water Policy, 14, 391–408. 80

Hall, P., Racine, J., Li, Q. 2011. Cross-validation and the estimation of conditional probability densities. Journal of the American Statistical Association. 94

Halla, M. 2012. Tax morale and compliance behavior: First evidence on a causal link. The BE Journal of Economic Analysis & Policy, 12. 3

Hanousek, J., Palda, F. 2004. Quality of government services and the civic duty to pay taxes in the Czech and Slovak Republics, and other transition countries. Kyklos, 57, 237–252. 4

Hashimzade, N., Myles, G. D., Page, F., Rablen, M. D. 2014. Social networks and occupational choice : The endogenous formation of attitudes and beliefs about tax compliance. Journal of Economic Psychology, 40, 134–146. 43

Hashimzade, N., Myles, G. D., Page, F., Rablen, M. D. 2015. The use of agent-based modelling to investigate tax compliance. Economics of Governance, 16, 143–164. 43

Hashimzade, N., Myles, G. D., Rablen, M. D. 2016. Predictive analytics and the targeting of audits. Journal of Economic Behavior and Organization, 124, 130–145. 43

Hindriks, J. 1999. On the incompatibility between revenue maximisation and tax progressivity. European Journal of Political Economy, 15, 123–140. 37

HM Treasury & HMRC. 2011. Tackling Tax Avoidance.Technical Report March, HM Treasury and HM Revenue and Customs. 6

HMRC. 2015. Measuring Tax Gaps 2015 Edition: Tax Gap Estimates for 2013-14. London, HM Revenue and Customs. 6, 35

Hsu, L. C. 2008. Experimental evidence on tax compliance and voluntary public good provision. National Tax Journal, 61, 205–223. 4

Huang, Y.-T., Pan, T.-C., Kao, J.-J. 2011. Performance assessment for municipal solid waste collection in taiwan. Journal of environmental management, 92, 1277–1283. 79

Internal Revenue Service. 2014. Internal Revenue Service Data Book 2013. Washington DC, IRS, 1–76. 21

ISTAT. 2011. Census of the general population. 84

Johnson, E. J., Shu, S. B., Dellaert, B. G., Fox, C., Goldstein, D. G., Häubl, G., Larrick, R. P., Payne, J. W., Peters, E., Schkade, D. et al. 2012. Beyond nudges: Tools of a choice architecture. Marketing Letters, 23, 487–504. 2, 10

Kahneman, D. 2003. Maps of bounded rationality: Psychology for behavioral economics. The American economic review, 93, 1449–1475. 2, 12

Kahneman, D., Tversky, A. 1979. Prospect theory: An analysis of decision under risk. Econometrica: Journal of the econometric society, 263–291. 27

Kalb, A. 2010. The impact of intergovernmental grants on cost efficiency: theory and evidence from german municipalities. Economic Analysis and Policy, 40, 23–48. 88

Kalb, A., Geys, B., Heinemann, F. 2012. Value for money? german local government efficiency in a comparative perspective. Applied economics, 44, 201–218. 80

Kalseth, J., Rattsø, J. 1995. Spending and overspending in local government administration: A minimum requirement approach applied to norway. European Journal of Political Economy, 11, 239–251. 79

Kim, Y. 2003. Income distribution and equilibrium multiplicity in a stigma-based model of tax evasion. Journal of Public Economics, 87, 1591–1616. 43

Kirchler, E., Hoelzl, E., Wahl, I. 2008. Enforced versus voluntary tax compliance: ”the slippery slope” framework. Journal of Economic Psychology, 29, 210–225. 3

Kirchler, E., Maciejovsky, B., Schneider, F. 2003. Everyday representations of tax avoidance, tax evasion, and tax flight: Do legal differences matter?. Journal of Economic Psychology, 24, 535–553. 13, 43

Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., Saez, E. 2011. Unwilling or unable to cheat? evidence from a tax audit experiment in denmark. Econometrica, 79, 651–692. 68

Kneip, A., Simar, L., Wilson, P. W. 2008. Asymptotics and consistent bootstraps for dea estimators in nonparametric frontier models. Econometric Theory, 24, 1663–1697. 98

Kneip, A., Simar, L., Wilson, P. W. 2011. A computationally efficient, consistent bootstrap for inference with non-parametric dea estimators. Computational Economics, 38, 483–515. 78, 98

Kneip, A., Simar, L., Wilson, P. W. 2015. When bias kills the variance: Central limit theorems for dea and fdh efficiency scores. Econometric Theory, 31, 394–422. 102

Kőszegi, B., Rabin, M. 2009. Reference-dependent consumption plans. The American Economic Review, 99, 909–936. 11

Kravis, I. B., Lipsey, R. E. 1982. The location of overseas production and production for export by U.S. multinational firms. Journal of International Economics, 12, 201–223. 78

Kristensen, O. P. 1983. Public versus private provision of governmental services: The case of danish fire protection services. Urban Studies, 20, 1–9. 79

Lang, O., Nöhrbaß, K.-H., Stahl, K. 1997. On income tax avoidance: the case of Germany. Journal of Public Economics, 66, 327–347. 10, 35

Langer, T., Weber, M. 2001. Prospect theory, mental accounting, and differences in aggregated and segregated evaluation of lottery portfolios. Management Science, 47, 716–733. 11

Lee, K. 2001. Tax evasion and self-insurance. Journal of Public Economics, 81, 73–81. 7, 43

Li, Q., Racine, J. S. 2008. Nonparametric estimation of conditional cdf and quantile functions with mixed categorical and continuous data. Journal of Business & Economic Statistics, 26, 423–434. 94

Li, Q., Racine, J. S. 2007. Nonparametric econometrics: theory and practice. Princeton University Press. 98, 101, 123

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. 80, 88, 127, 128

Lockwood, B., Porcelli, F. 2013. Incentive Schemes for Local Government: Theory and Evidence from Comprehensive Performance Assessment in England. American Economic Journal: Economic Policy, 5, p. 40. 76

Loikkanen, H., Susiluoto, I. et al. 2005. Cost efficiency of finnish municipalities in basic service provision 1994-2002. Urban Public Economics Review, 4, 39–64. 78, 80

Marhuenda, F., Ortuño-Ortn, I. 1994. Honesty vs. progressiveness in income tax enforcement problems. 37, 45

McCaffery, E. J., Baron, J. 2004. Framing and taxation: Evaluation of tax policies involving household composition. Journal of Economic Psychology, 25, 679–705. 12

Mookherjee, D., Png, I. 1989. Optimal auditing, insurance, and redistribution. The Quarterly Journal of Economics, 104, 399–415. 37

Morton, S. 1993. Strategic auditing for fraud. The Accounting Review, 68, p. 825. 36

Myerson, R. B. 1979. Incentive compatibility and the bargaining problem. Econometrica, 47, 61–74. 45

Myerson, R. B. 1982. Optimal coordination mechanisms in generalized principal-agent problems. Journal of Mathematical Economics, 10, 67–81. 45

Myerson, R. B. 1989. Mechanism design. In J. Eatwell, M. Milgate, P. Newman eds. Allocation, Information and Markets, New York, McMillan Press, 191–206. 45

Myles, G. D., Naylor, R. A. 1996. A model of tax evasion with group conformity and social customs. European Journal of Political Economy, 12, 49–66. 3, 43

Nijkamp, P., 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, 249–274. 80

Oates, W. E. 1999. An Essay on Fiscal Federalism. Journal of Economic Literature, 37, 1120–1149. 76

Oates, W. E. 2005. Toward A Second-Generation Theory of Fiscal Federalism. International Tax and Public Finance, 349–373. 76, 82

Pagan, A., Ullah, A. 1999. Nonparametric econometrics. Cambridge university press. 101

Papke, L. E. 1991. Interstate business new firm location tax differentials and evidence from panel data. Journal of Public Economics, 45, 47–68. 77

Patel, P., Vega, M. 1999. Patterns of internationalisation of corporate technology: location vs. home country advantages. Research Policy, 28, 145–155. 77

Persson, T., Roland, G., Tabellini, G. 1997. Separation of powers and political accountability. The Quarterly Journal of Economics, 112, 1163–1202. 77

Persson, T., Tabellini, G. 1996. Federal fiscal constitutions: risk sharing and moral hazard. Econometrica, 3, 623–646. 82

Picazo-Tadeo, A. J., Gonzalez-Gomez, F., Saez-Fernandez, F. J. 2009. Accounting for operating environments in measuring water utilities managerial efficiency. The Service Industries Journal, 29, 761–773. 80

Piketty, T., Saez, E., Stantcheva, S. 2014. Optimal taxation of top labor incomes: A tale of three elasticities. American economic journal: economic policy, 6, 230–271. 9, 40

Pina, V., Torres, L. 2001. Analysis of the efficiency of local government services delivery. an application to urban public transport. Transportation Research Part A: Policy and Practice, 35, 929–944. 80

Piolatto, A., Rablen, M. D. 2016. Prospect Theory and Tax Evasion: A Reconsideration of the Yitzhaki Puzzle. Theory and Decision. 61

Politis, D. N., Romano, J. P., Wolf, M. 2001. On the asymptotic theory of subsampling. Statistica Sinica, 1105–1124. 98, 100

Prebble, Z., Prebble, J. 2010. Morality of tax avoidance, the. Creighton University School of Law, 18, 693–745. 12

Prieto, A. M., ZofIo, J. L. 2001. Evaluating effectiveness in public provision of infrastructure and equipment: the case of spanish municipalities. Journal of productivity Analysis, 15, 41–58. 79

Prud'homme, R. 1995. The dangers of decentralization. The World Bank Research Observer, 10, 201–220. 76

Rabin, M., Weizsäcker, G. 2009. Narrow bracketing and dominated choices. The American economic review, 99, 1508–1543. 1, 7, 11

Rablen, M. D. 2014. Audit probability versus effectiveness: The Beckerian approach revisited. Journal of Public Economic Theory, 16, 322–342. 30, 70

Read, D., Antonides, G., Van den Ouden, L., Trienekens, H. 2001. Which is better: simultaneous or sequential choice? Organizational behavior and human decision processes, 84, 54–70. 11

Read, D., Loewenstein, G. F., Rabin, M. 1999. Choice bracketing. Journal of risk and uncertainty : JRU, 19, 171–197. 11

Reinganum, J. F., Wilde, L. L. 1985. Income tax compliance in a principal-agent framework. Journal of Public Economics, 26. 36

Reinganum, J. F., Wilde, L. L. 1986. Equilibrium verification and reporting policies in a model of tax compliance. International Economic Review, 27, 739–760. 42

Reinganum, J. F., Wilde, L. L. 1991. Equilibrium enforcement and compliance in the presence of tax practitioners. Journal of Law, Economics, & Organization, 7, 163–181. 14, 41

Revelli, F., Tovmo, P. 2007. Revealed yardstick competition: Local government efficiency patterns in norway. Journal of Urban Economics, 62, 121–134. 127

Rodden, J. 2002. The dilemma of fiscal federalism: Hard and soft budget constraints around the world. 82

Rogge, N., De Jaeger, S. 2012. Evaluating the efficiency of municipalities in collecting and processing municipal solid waste: A shared input dea-model. Waste management, 32, 1968–1978. 79

Rogge, N., De Jaeger, S. 2013. Measuring and explaining the cost efficiency of municipal solid waste collection and processing services. Omega, 41, 653–664. 102

Sanchez, I., Sobel, J. 1993. Hierarchical design and enforcement of income tax policies. Journal of Public Economics, 50, 345–369. 36

Sarkis, J., Dijkshoorn, J. 2007. Relationships between solid waste management performance and environmental practice adoption in welsh small and medium-sized enterprises (smes). International Journal of Production Research, 45, 4989–5015. 79

Schmidt, S. S. 2008. The measurement of productive efficiency and productivity growth. Oxford University Press. 88

Scholz, J. T., Lubell, M. 1998. Trust and taxpaying: Testing the heuristic approach to collective action. American Journal of Political Science, 398–417. 3

Schwartz, R. D., Orleans, S. 1967. On legal sanctions. The University of Chicago Law Review, 34, 274–300. 3

Scotchmer, S. 1987. Audit classes and tax enforcement policy. American Economic Review, 77, 229–233. 36

Seabright, P. 1995. Accountability and Decentralization in Government: An Incomplete Contracts Model. European Economic Review, 40, 61–89. 77

Settimi, C., Vidoli, F., Fusco, E., Ballanti, D. 2014. Estimating technical efficiency in the italian municipalities. Procedia Economics and Finance, 17, 131–137. 79

Shah, A. 2012. Public services and expenditure need equalization: reflections on principles and worldwide comparative practices. World Bank Policy Research Working Paper. 76

Shepherd, R. W. 1970. Theory of cost and Production Function. Princeton University Press. 92

Sikka, P. 2012. The tax avoidance industry. Radical statistics, 107, 15–30. 8

Simar, L. 1996. Aspects of statistical analysis in DEA-type frontier models. Journal of Productivity Analysis, 185, 177–185. 104, 129

Simar, L. 2003. Detecting Outliers in Frontiers Models: A Simple Approach. Journal of Productivity Analysis, 20, 391–424. 104

Simar, L., Wilson, P. W. 2007. Estimation and inference in two-stage, semi-parametric models of production processes. Journal of econometrics, 136, 31–64. 78, 80, 100

Simar, L., Wilson, P. W. 2011a. Inference by the m out of n bootstrap in nonparametric frontier models. Journal of Productivity Analysis, 36, 33–53. 98

Simar, L., Wilson, P. W. 2011b. Two-stage dea: caveat emptor. Journal of Productivity Analysis, 36, 205–218. 78, 79, 100

Slemrod, J. 1985. An empirical test for tax evasion. The Review of Economics and Statistics, 67, 232–238. 68

Slemrod, J. 1995. Income creation or income shifting? behavioral responses to the tax reform act of 1986. The American Economic Review, 85, 175–180. 9, 40

Slemrod, J. 1996. High-income families and the tax changes of the 1980s: The anatomy of behavioral response. In M. Feldstein, J. M. Poterba eds. Empirical foundations of household taxation, Chicago, University of Chicago Press, 169–192. 40

Slemrod, J., Kopczuk, W. 2002. The optimal elasticity of taxable income. Journal of Public Economics, 84, 91–112. 9, 40

Slemrod, J., Yitzhaki, S. 2002. Tax avoidance, evasion and administration. In A. J. Auerbach, M. Feldstein eds. Handbook of Public Economics, 1, Amsterdam, North-Holland, 1423–1470. 9, 34

Snow, A., Warren, R. S. 2005a. Tax evasion under random audits with uncertain detection. Economics Letters, 88, 97–100. 70

Snow, A., Warren, R. S. 2005b. Ambiguity about audit probability, tax compliance, and taxpayer welfare. Economic Inquiry, 43, 865–871. 30, 70

Srinivasan, T. N. 1973. Tax evasion: A model. Journal of public economics, 2, 339–346. 35

Stevens, P. 2013. The changing limits of acceptable tax avoidance. Journal of public economics, 1152, 17–21. 14

Stiglitz, J. E. 1986. The general theory of tax avoidance. National tax journal, 38, 325–338. 9, 34

Stone, J. 2015. Most people think legal tax avoidance is just as wrong as illegal tax evasion, poll suggests. 43

Torgler, B. 2002. Speaking to theorists and searching for facts: Tax morale and tax compliance in experiments. Journal of Economic Surveys, 16, 657–683. 3

Torgler, B., Schneider, F. 2009. The impact of tax morale and institutional quality on the shadow economy. Journal of Economic Psychology, 30, 228–245. 3

Traxler, C. 2010. Social norms and conditional cooperative taxpayers. European Journal of Political Economy, 26, 89–103. 3

Tversky, A., Kahneman, D. 1981. The framing of decisions and the psychology of choice. Science, 211, 453–458. 11

Uribe-Teran, C. 2015. Down the rabbit–hole: Tax avoidance and the elasticity of taxable income. Mimeo. 40

Verschelde, M., 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, 214–225. 80

Wallschutzky, I. G. 1984. Possible causes of tax evasion. Journal of Economic Psychology, 5, 371–384. 69

Wheeler, D., Mody, A. 1992. International decisions investment location. Journal of International Economics, 33, 57–76. 78

Worthington, A., Dollery, B. 2000. Measuring efficiency in local governments' planning and regulatory function. Public Productivity & Management Review, 469–485. 84

Worthington, A., Dollery, B. 2001. Measuring efficiency in local government: an analysis of new south wales municipalities' domestic waste management function. Policy Studies Journal, 29, 232–249. 79

Worthington, A. C. 2000. Cost efficiency in australian local government: A comparative analysis of mathematical programming and econometrical approaches. Financial Accountability & Management, 16, 201–223. 80, 127

Yitzhaki, S. 1974. Income tax evasion: A theoretical analysis. Journal of public economics, 3, 201–202. 6, 20, 24, 35, 39, 51, 61