ARTICLE IN PRESS

Economic Systems xxx (xxxx) xxx

Contents lists available at ScienceDirect

Economic Systems

journal homepage: www.elsevier.com/locate/ecosys

Audit, presumptive taxation and efficiency: An integrated approach for tax compliance analysis

Giancarlo Ferrara[a], Valeria Bucci[a,b,*], Arianna Campagna[a]

[a] *Soluzioni per il Sistema Economico SpA, Rome, Italy*
[b] *Cassa Depositi e Prestiti SpA, Rome, Italy*

A R T I C L E   I N F O

JEL Classification:
H26
H32
C14
D24

Keywords:
Tax compliance
Presumptive taxation
Efficiency
Stochastic frontier
Business sector studies

A B S T R A C T

Audit cut-off rules are policy tools widely adopted by fiscal authorities with the aim to improve voluntary tax compliance. Despite their usefulness, audits are costly for the fiscal authority, so determining how best to allocate controls is a key policy issue. In the audit framework presumptive taxation could be used to identify the cut-off rule, allowing fiscal authorities to uncover firms' under-reporting. However, an audit rule based on presumptions does not allow distinguishing whether the presence of under-reporting is due to voluntary non-compliance or to a lack of managerial skills. Therefore, this paper proposes to combine an audit rule based on a presumptive cut-off with a measure of efficiency, with the aim to solve the main weakness of the presumptive methods. In particular, we develop an integrated approach that is able to support the audit activities of fiscal authorities. For illustrative purposes we support our approach with an empirical application based on a sample of Italian firms.

1. Introduction

Increasing voluntary tax compliance in order to fight tax evasion represents a critical issue for many governments worldwide. Several motivations could explain this need. The most common reason is that tax evasion causes significant revenue losses, implying reductions in public expenditures and cuts in public services. The misallocation of resources is another consequence of tax evasion that could force economic agents to alter their behaviour, for example by modifying the labour supply or investment schedules (Alm, 2019). Tax evasion could also prevent the socially optimal redistribution of resources by altering the distribution of income unfairly. Furthermore, a fiscal system affected by tax evasion characterised by unfair behaviour and non-observance of the law could undermine citizens' trust in their government.

The economics-of-crime literature, stemming from the study by Allinhgam and Sadmo (1972), identifies audits as a useful instrument in coping with tax evasion.[1] Audit activity could raise tax revenues through the assessment of additional taxes, interests and penalties on audited taxpayers, the so-called direct effect. Furthermore, the probability of detection and penalties could deter future noncompliance among both audited and unaudited taxpayers – the indirect effects (Kasper and Alm, 2022). However, despite their usefulness, tax audits are costly for the fiscal authority, so determining the number of audits and how best to allocate controls are key policy issues (Slemrod and Yitzhaki, 2002).

* Corresponding author at: Soluzioni per il Sistema Economico SpA, Rome, Italy.
E-mail address: valeriabucci@gmail.com (V. Bucci).

[1] For a thorough review on tax enforcement instruments, see Slemrod (2019).

https://doi.org/10.1016/j.ecosys.2023.101099
Received 31 January 2022; Received in revised form 11 November 2022; Accepted 14 November 2022
Available online xxxx

Tax auditing schemes range from random to the cut-off audit rule.[2] The first randomly selects taxpayers to be audited, while the cut-off scheme involves a threshold identifying as auditable only taxpayers reporting an income lower than a certain value (Arachi and Santoro, 2007). The empirical literature shows that cut-off rules lead to higher tax compliance than random audits (Alm et al., 1993; Kirchler, 2007; Reinganum and Wilde, 1985). However, the limited resources for audits usually imply more frequent inspections for lower-income classes, in order to prevent their imitation by high-income taxpayers. Therefore, the definition of the cut-off rule represents a crucial element for the optimal audit scheme maximising tax revenues.

A useful instrument to support tax auditing could be provided by presumptive taxation. As an alternative to the regular methods used to compute actual taxable income using taxpayers' accounts, presumptive taxation methods reconstruct taxpayers' income through administrative practices based on supplementary information. Such information refers to variables not included in the standard computation of taxable income but linked to income generation and easily achievable by tax authorities.

In the audit framework, presumptive values could be used to set the cut-off rule: exclusively taxpayers reporting values lower than those reconstructed through the presumptive methods are auditable by the fiscal authority. Therefore, from the fiscal authority's point of view, the adoption of presumptive methods could significantly reduce the compliance costs by overcoming administrative weaknesses to ascertain the actual tax base in the case of poor record-keeping, avoiding the costs linked to conventional assessment procedures.[3]

However, presumptive methods face several limits. Usually, the taxpayers know the presumptive rules perfectly, and consequently there is a strong incentive to manipulate reported revenues and costs in order to reach the presumed value. Another weakness is that usually the presumptive approach adopts a regression framework to model the behaviour of the average firm. This average setting could induce firms above the presumed values to reduce their reported numbers, hiding their extra performance from the tax authority. Finally, presumptive taxation methods allow authorities to discover firms' under-reporting (i.e., presumptive values are higher than the data reported by taxpayers) without discerning the motivations for such discrepancies. Usually, presumptive regimes neglect inefficiency, not considering whether a firm with particular characteristics (i.e. type, size, location, number of employees, etc.) sustains unusually high business expenses and/or gains unusually low profits in a given year (Logue and Vettori, 2011). Therefore, as stated by Yitzhaki (2007), presumptive taxation systems imply that inefficient taxpayers pay more taxes.

This paper starts by considering the potential advantages of an audit framework using presumptive methods to define the cut-off rule, in particular in terms of cost reduction, and proposes to combine this approach with a measure of efficiency with the aim to overcome the main weakness of these methods. In particular, we propose to estimate a production function using stochastic frontier analysis based on a generalised additive model specification (Vidoli and Ferrara, 2015), setting the threshold at the level of the most efficient firm (Kumbhakar and Lovell, 2000). We solve a critical issue of presumptive taxation, represented by the presence of under-reporting (accidental or not) of inputs, by adopting a two-step methodology to estimate the production inputs potentially affected by under-reporting. Our integrated approach reduces the unfair behaviours generated by both the consideration of the average firm and the advantage of revenue/cost manipulation, i.e. taxpayers' learning-by-doing attitude. Therefore, we obtain an integrated approach for tax compliance analysis that is able to support the fiscal authority's audit activity, allowing to distinguish situations of under-reporting ascribable to voluntary non-compliant behaviour from those due to a lack of managerial skills, potentially implying taxpayers' inefficiency.

This paper contributes to the existing literature in several ways. First, we contribute to the literature on audits and presumptions by pointing out that, if integrated with efficiency measures, such methods could provide a useful instrument to support the fiscal authority in identifying taxpayers adopting potentially anomalous behaviours. To the best of our knowledge, very few studies have analysed audit instruments using presumptive methods to define the threshold value, failing to find convincing evidence for their effectiveness in increasing voluntary tax compliance.[4] In particular, an instrument combining audit and presumptive taxation that received some consideration in the empirical literature is provided by the Italian *Business Sector Studies (BSS)*. The empirical findings show that taxpayers tend to manipulate reported values in order to reduce the presumed turnover rather than voluntarily becoming more compliant (Bucci, 2020; Pisani, 2004; Santoro, 2008; Santoro and Fiorio, 2011). Second, we show that the concept of firm efficiency is suitable for application in tax compliance analysis. The existing literature has long investigated the determinants of firms' noncompliance (Alvarez and Crespi, 2003; Bottasso and Sembenelli, 2004; Heshmati, 2003), but the role of efficiency in firms' choices has been rather neglected. While it is possible to find scant evidence on the role of firm efficiency for the corporate capital structure (Berger and Bonaccorsi di Patti, 2006; Margaritis and Psillaki, 2007), to the best of our knowledge this is the first paper that takes the effect of taxpayers' efficiency into consideration in tax compliance analysis. Third, we develop a measure of tax compliance that could allow us to identify whether some taxpayer characteristics might affect non-compliant behaviours (Slemrod, 2007). For example, analysing taxpayers belonging to different economic sectors would allow us to identify different compliance patterns for firms within different industries (Gokalp et al., 2017; Tedds, 2010). Finally, for illustrative purposes, we provide an empirical application supporting the potentiality of our approach. The empirical application has been developed on the basis of the Italian *BSS*, but can be adapted for different audit methods using presumptions to define the cut-off rule.

The remainder of this paper is organised as follows. In Section 2, we develop and discuss the empirical model. In Section 3, we describe the Italian *BSS* framework and present the main findings of the empirical application. The final section offers some concluding remarks and discusses the main issues for future research.

[2] Reinganum and Wilde (1985) provide a theoretical description of these audit rules.

[3] Presumptive taxation methods may be very useful when the books and records are difficult to verify for the tax authorities or when such values do not correctly reflect the taxable capacity (Martins and Sa, 2018; Slemrod and Yitzhaki, 1994).

[4] For a comprehensive survey of the empirical literature on the effectiveness of presumptive methods see Bucci (2020).

2. The empirical model: the integrated approach

2.1. The efficiency score estimation

This paper proposes an integrated approach for tax compliance analysis combining audit, presumptive taxation and efficiency.

A critical issue of presumptive taxation methods is represented by the under-reporting (accidental or not) of inputs by firms. Therefore, with the aim of solving this issue, we propose a two-step procedure. In particular, in the first step we estimate a linear regression model for each input k potentially affected by under-reporting x_k, using an input-specific vector of explanatory variables $\boldsymbol{Z}_k$:

$$x_{ik} = \alpha + \boldsymbol{\gamma Z}_{ik} + \in_{ik} \text{ for each firm } i = 1, \ldots, n \tag{1}$$

A general assumption of presumptive taxation methods is to presume taxpayer's income through information not considered in the standard computation of taxable income and to focus on taxpayers declaring values lower than the presumed ones. Therefore, we apply the same general concept when investigating the input declaration, focusing only on inputs potentially affected by under-reporting. When the input reported by firm i (x_{ik}) is lower than the fitted value given by the corresponding expected value (i.e. $\hat{x}_{ik} = E(\hat{x}_k)$), then the fitted value is used in the production function estimation, according to the following formula:

$$x_{ik}^* = \max(x_{ik}, \hat{x}_{ik}) i = 1, \ldots, n. \tag{2}$$

We also consider the possibility of setting different reference (i.e. minimum) values for $\hat{x}_{ik}$. More specifically, in order to evaluate the appropriateness of our choice to use the fitted values, for $\hat{x}_k$ we alternatively consider $E(\hat{x}_k)$, $E(\hat{x}_k) - \hat{\sigma}$ and $E(\hat{x}_k) - 2 \times \hat{\sigma}$, where σ is the standard error of residuals obtained by the regression model.

The second step proceeds with the estimation of the production function parameters. Our starting point is the parametric stochastic frontier model introduced by Aigner et al. (1977) and Meeusen and van den Broeck (1977).[5] In particular, we define the production function as a combination of a response function and a composite error term:

$$y_i = f(\boldsymbol{x_i}, \beta) + v_i - u_i i = 1, \ldots, n, \tag{3}$$

where $f(\cdot)$ represents the production relationship between the vector of inputs ($\boldsymbol{x}_i$), potentially resulting from the combination of Eqs. (1) and (2), and the maximum level of output achievable by firm i (y_i), and $\boldsymbol{\beta}$ is the vector of parameters to be estimated through the stochastic frontier model. The composite error term is given by the term v_i, which represents symmetric disturbance and is assumed to be $i.\ i.\ d.$ ($N(0, \sigma_v^2)$), and the term u_i, which represents the error component reflecting technical inefficiency and is assumed to be distributed independently of v_i as a half-normal truncated distribution above zero ($N^+(0, \sigma_u^2)$).

In spite of its simple computation and interpretation, this model does not allow for too much flexibility. Indeed, forcing $f(\cdot)$ to belong to a fully parametric family of functions (i.e., translog or Cobb-Douglas) might lead to a non-negligible bias in the model specification and to misleading conclusions about the link between inputs and output. To overcome drawbacks due to the specification of a particular production function, Fan et al. (1996) introduce a two-step pseudo-likelihood procedure to estimate the stochastic frontier model where the functional form of the frontier is unknown and estimated via kernel regression.[6] In this work, we consider the Vidoli and Ferrara (2015) model that extends Fan et al.'s (1996) approach by specifying a Generalised Additive Model framework (GAM). In a regression context with normal response, the model can be expressed as:

$$E(Y \mid \boldsymbol{X} = \boldsymbol{x}) = \psi_0 + \sum_{j=1}^{p} \psi_j(X_j), \tag{4}$$

where the $\psi_j(\cdot)'s$ are smooth functions standardised so that $E[\psi_j(X_j)] = 0$ (Hastie and Tibshirani, 1990). This model takes into account the variability of the response through an additive function of the inputs, as in the corresponding stochastic frontier model.

The main advantages of using the GAM specification for the stochastic frontier analysis (GAM-SFA) over the standard approaches are: a) the consideration that the transformations ψ_j's are determined simultaneously and the non-linear fits can potentially make more accurate predictions of the response; b) the non-parametric estimators of the unknown functions ψ_j are able to avoid the curse of dimensionality since each additive terms is estimated using a univariate smoother; c) the marginal response function ψ_j shows how the prediction changes with respect to X_j, as in an additive liner model. Furthermore, the gradients of the non-parametric model can be interpreted as partial output elasticities and their sum (i.e. elasticity of scale: eos) highlights useful information about specific return to scale.

In a cross-sectional setting, model (3) becomes:

$$y_i = \Psi(\boldsymbol{x}_i) + v_i - u_i i = 1, \ldots, n, \tag{5}$$

where the unknown function $\Psi(\cdot)$ is modelled via GAM (4). When the response is measured in logs, in frontier models the relative estimates of the technical efficiency for each unit is obtained by.[7].

---

[5] In the literature different stochastic frontier models have been proposed. Overviews of developments in this area have been provided by Sickles and Zelenyuk (2019) and Kumbhakar and Lovell (2000).

[6] More recently, Kumbhakar et al. (2007) proposed a new approach based on the local maximum likelihood principle.

[7] The model estimation has been carried out using the R Environment (R Core Team, 2020 using the semsfa and mgcv packages. In particular, the $f_j's$ smooth functions are represented using thin plate regression splines avoiding the knot placement problems of conventional regression spline modelling (Wood, 2003). Please see Wood (2006) for the test statistics related to smooth terms and the graphical representations for the analysis and interpretation of the ψ_j's. For a recent review of the stochastic frontier analysis using R see Ferrara (2020).

$$TE_i = \exp\{-\hat{u}_i\}. \tag{6}$$

2.2. Integrated analysis

Our integrated approach has been developed considering an audit framework based on a presumptive method. In particular, the presumptive method reconstructs taxpayers' income through administrative practices based on information supplementary to accounting and fiscal data, linked to income generation and easily achievable by the tax authority. Then, in the audit framework, the presumed values are used to identify the cut-off rule: taxpayers reporting values lower than the presumptions are auditable by the fiscal authority.

We propose to combine audits and presumptions with the efficiency scores estimated for each firm. In particular, we consider the efficiency score and the relative distance between the reported and presumed turnover simultaneously for each firm, i.e. the auditable taxpayers.[8] We assume that in the presence of a negative gap between the recorded and presumed values associated with a low efficiency score, it is more likely that under-reporting is due to firm's inefficiency rather than to voluntary non-compliance. Otherwise, voluntary non-compliant behaviour is more likely in the presence of highly efficient firms exhibiting negative differences between their reported and presumed values.

The integrated approach can be represented through a graph analysis, like Fig. 1, where each taxpayer is represented in terms of their efficiency score[9] and their presumptive result (i.e., the difference between reported and presumed turnover). Therefore, by combining these dimensions, we are able to categorise taxpayers into 4 different quadrants: the firms in quadrants I and II report a turnover higher than the value presumed by the fiscal authority. Therefore, such taxpayers are exempt from the fiscal authority's audit since it is less likely that they are adopting voluntary non-compliant behaviours.[10]

In contrast, firms in quadrants III and IV have negative gaps between their declared and presumed turnover. In a standard audit framework based on a presumptive cut-off rule, taxpayers in both quadrants III and IV should be auditable by the fiscal authority. However, there is a discriminant between quadrants III and IV, relying on the efficiency score. The definition of a threshold value for the efficiency score would allow distinguishing the group of high efficiency firms (quadrant III) from the low efficiency ones (quadrant IV).[11]

Therefore, while for quadrant IV the negative gap between reported and presumed turnover is linked to a lower efficiency score, the firms in quadrant III do not reach the presumed values, despite a high level of efficiency. Therefore, the integrated approach allows us to identify quadrant III as the area where there is more likely to be a mass of firms that adopt potentially anomalous behaviours. Thus the firms in quadrant III could require additional investigation, identifying the group of taxpayers to be audited with priority by the fiscal authority, and defining a ranking for controls.

2.3. Tax compliance analysis

Further, we develop a measure of tax compliance analysis that could provide evidence for the validity of our integrated methodology.

The empirical analysis of tax compliance is notoriously difficult due to the lack of reliable information on taxpayers' compliance, as it is complicated to define and measure (Alm, 2019; Slemrod and Weber, 2012; Torgler, 2016). We address this issue by assuming that a difference between the value of the presumed and reported turnover is due to taxpayers' non-compliance, implying a difference between the potential taxes collected and the amount of taxes actually paid.[12]

Therefore, we compute a simple non-compliance measure (*SSR*) based on the sum of squared residuals of the deviation of a firm's turnover from the presumed threshold. Further, we define a more stringent measure for non-compliance (*NW-SSR*), weighting the *SSR* with firms' efficiency scores, taking into account exclusively firms below the presumed value that need further investigation by the tax authority. A higher *NW-SSR* implies lower tax compliance.

The tax compliance analysis would allow us to compare and rank different groups of taxpayers.[13]

3. Empirical application

For illustrative purposes, in order to test the potentiality of our integrated approach, we implement an empirical application based on the Italian *Business Sector Studies (BSS)*, an example for an audit instrument using the presumptive method to define the cut-off rule.

The *Business Sector Studies* estimate the value of the turnover of SMEs and self-employed using the information declared by these taxpayers.[14] The *BSS* is based on the hypothesis that two firms produce the same turnover and should have to pay the same amount of

[8] In what follows, the relative distance is computed as the difference between recorded and presumed turnover as a share of the presumed value.

[9] We consider all potential values the efficiency score could assume, ranging from 0 for the least efficient firm to 1 for the most efficient firm.

[10] In the spirit of audit based on a cut-off rule, we excluded from the group of potentially auditable firms those reporting a turnover higher than the presumed value. However, it is possible that firms could reach the presumed turnover due to manipulating the reported variables.

[11] In this graph analysis, for illustrative purposes, we set the threshold value equal to 0.5, but in the empirical analysis, it is possible to define the threshold value using different measures, such as the first quartile of the corresponding distribution, or the value of a different percentile.

[12] The idea behind this is that the higher the deviation of each taxpayer from the presumed value, the higher the value of the under-reporting of turnover and, consequently, the lower the tax compliance.

[13] It is important to highlight that a simple comparison between *NW-SSR* measures could result in biased conclusions due to the different dimensions of taxpayer groups. Therefore, we suggest weighting such measures according to the number of firms below the presumed turnover value before comparing them.

[14] For a detailed description of the *BSS* structure, see Fiorio et al. (2013).

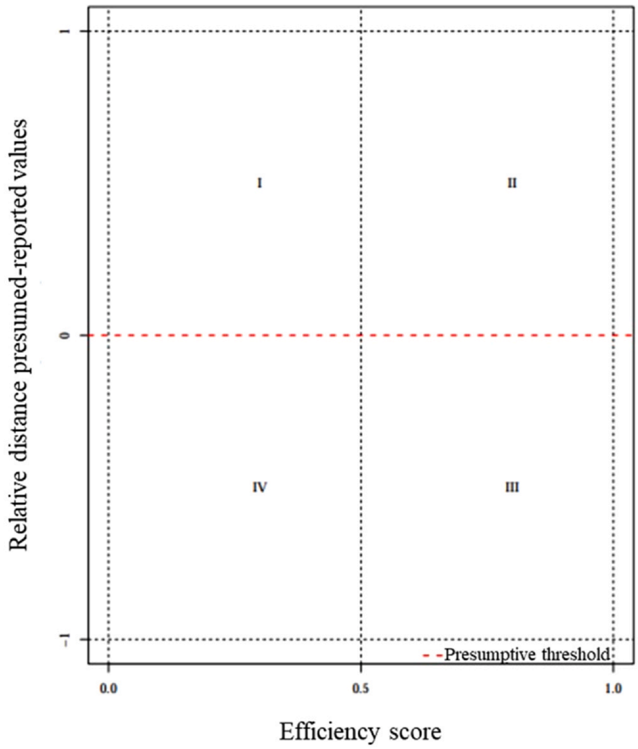

Fig. 1. Integrated analysis.

taxes if they belong to the same economic sector and face the same combination of input costs (i.e., the two firms share the same production function). Therefore, given that the "true" level of turnover is unknown, an econometric regression model is used to estimate the presumptive value, conditional on taxpayers' characteristics, obtaining a vector of coefficients showing the presumptive productivity of each input.

The reported turnover is then compared to the presumed value, allowing authorities to uncover situations of under-reporting of revenues and costs. In the case of discrepancies between reported and presumed values, firms defined as "non-congruous" can adapt to the presumed results or appeal to the fiscal authority by providing evidence that their actual performance is lower than the presumed one.

When the presumptive regime is too simple and too favourable, firms could choose to alter their behaviour by manipulating reporting data in order to rely on the presumptive regime because of the opportunity to reduce their tax burden (Dube, 2018; Memon, 2013).

An extremely important element of the *BSS* framework concerns the timing. Usually, when firms report revenues and costs, the rules used by the tax authority to estimate presumptive values are known. Consequently, it is possible for firms to modify the reported numbers in order to reach the presumed values. Another weakness of the presumptive *BSS* system is represented by the average setting: the presumptive values are estimated using an OLS regression to catch the average firm's behaviour. This average setting could be exploited to under-report revenues for firms above the presumed values, inducing the firms to hide their extra performance from the tax authority. Pisani (2004) shows that over the years a growing number of taxpayers declared revenues around the mean value of the distribution, highlighting a process of learning-by-doing by firms.

The important main limit of the *BSS* framework is the inability to identify the motivations for the presence of under-reporting (i.e., presumptive values higher than reported numbers). These methods do not allow disentangling whether the under-reporting is ascribable to voluntary non-compliance or to a lack of managerial skills, potentially implying firms' inefficiency.

3.1. Data description

The empirical application is based on a cross-sectional dataset of firms active in the retail and services sectors, subject to the *Business Sector Studies* in the fiscal year 2006. In particular, we used a sample composed of 222 companies belonging to two different sectors: retail (102 firms) and services (120 firms).[15]

Table 1 provides the descriptive statistics of the main variables used in the empirical analysis: the number of employees, capital stock, and input costs, which were divided into the costs for goods and costs for services. Additionally, information concerning the amount of gas used in cubic metres is provided only for the services sample (Panel (b) of Table 1).

[15] We consider the same cross-sectional dataset concerning the *BSS* regime already analyzed in Ferrara (2011).

Table 1
Summary statistics.

(a) Retail Sector				
Variables	**Sample (n = 102)**		**BSS (n = 10,235)**	
	Mean	**St. Dev.**	**Mean**	**St. Dev.**
Turnover	230	143	203	310
Capital stock	93	79	83	125
Number of employees	2	1	2	5
Costs of materials	116	78	108	183
Costs of services	23	23	19	39
(b) Services Sector				
Variables	**Sample (n = 120)**		**BSS (n = 88,110)**	
	Mean	**St. Dev.**	**Mean**	**St. Dev.**
Turnover	177	169	189	259
Capital stock	104	154	77	129
Number of employees	3	1	3	3
Costs of materials	65	61	78	104
Costs of services	19	21	22	41
Gas consumption	1124	2201	3500	7700

Note: All variables are expressed in thousands of euros, while gas consumption is expressed in cubic metres.

By comparing each sample with the corresponding *BSS* values, we observe that, although the two samples have not been randomly selected, their average values are rather close to the ones of the corresponding *BSS*. The confidence intervals associated with each variable considered in the analysis overlap substantially for both the samples and the BSS.

By comparing the values by economic sector, it emerges that the retail sample reports on average higher turnover and higher costs in terms of both goods and services purchased. Further, the retail sample employs fewer employees and records lower capital stock. Finally, taxpayers belonging to the services sample use more than one thousand cubic metres of gas.

3.2. The retail sector

The empirical application for the estimation of the efficiency scores starts with the regression analysis.

As illustrated in 2.1, in order to avoid the problem of under-reporting (accidental or not) of inputs by taxpayers, we adopt a two-step procedure. In the first step, we estimate a linear regression model for the labour variable, the input potentially mainly affected by under-reporting.[16] We specify the labour equation as follows:

$$L_i = \gamma_1 K_i + \gamma_2 M_i + \gamma_3 S_i + \nu_i i = 1, \ldots, n, \tag{7}$$

where L_i is the number of employees employed by company i, K_i is the capital stock, S_i is the costs for services, and M_i is the costs for materials (all the variables are expressed in logs).

The second step proceeds with the estimation of the production function by implementing the stochastic frontier analysis based on the GAM-SFA model:

$$y_i = \psi(\beta_0 + \beta_1 K_i + \beta_2 L_i^* + \beta_3 M_i) + v_i - u_i i = 1, \ldots, n, \tag{8}$$

where y_i is the log value of the turnover reported by company i and L_i^* is the higher number between the reported and the fitted value for the labour variable obtained following Eqs. (1) and (2).

In Fig. 2 we show the estimated marginal effects of the GAM frontier, providing evidence of non-linearity, in contrast with the corresponding Cobb-Douglas specification of the production function, linear or logs.[17]

Fig. 3 shows that the efficiency scores estimated for retail range from 0.66 to 0.96 and that a high share of firms in this sample is quite efficient, since most of the firms report an efficiency score higher than 0.9 (approximately 54%). Given that the nonparametric specification allows for subject specific partial elasticities, we also report the distribution of the elasticity of scale, given by the sum of the partial elasticities, highlighting increasing return to scale for all firms.

The efficiency scores estimated through the GAM-SFA model for each firm in the sample are then combined with the relative distances between the reported turnover and the presumptive values estimated by the Italian tax authority under the *BSS* regime: the *BSS* congruity threshold. The congruity analysis shows that taxpayers are both above and below the *BSS* threshold, with some predominance above. However, when we jointly consider both dimensions (efficiency and congruity), some interesting evidence emerges. The low efficiency firms (i.e., efficiency score below 0.9) are quite uniformly spread out along the estimated threshold. For the high efficiency scores (i.e., higher than 0.9), the number of firms recording turnover above the estimated threshold is higher than the share of firms with a reported turnover below the *BSS* threshold (54% vs.

[16] As our empirical analysis is exclusively illustrative, we considered only the labour input as under-reported, for the sake of simplicity. However, it would be possible to further extend this analysis by also considering additional inputs potentially affected by under-reporting.

[17] The estimation results are reported in the Appendix.

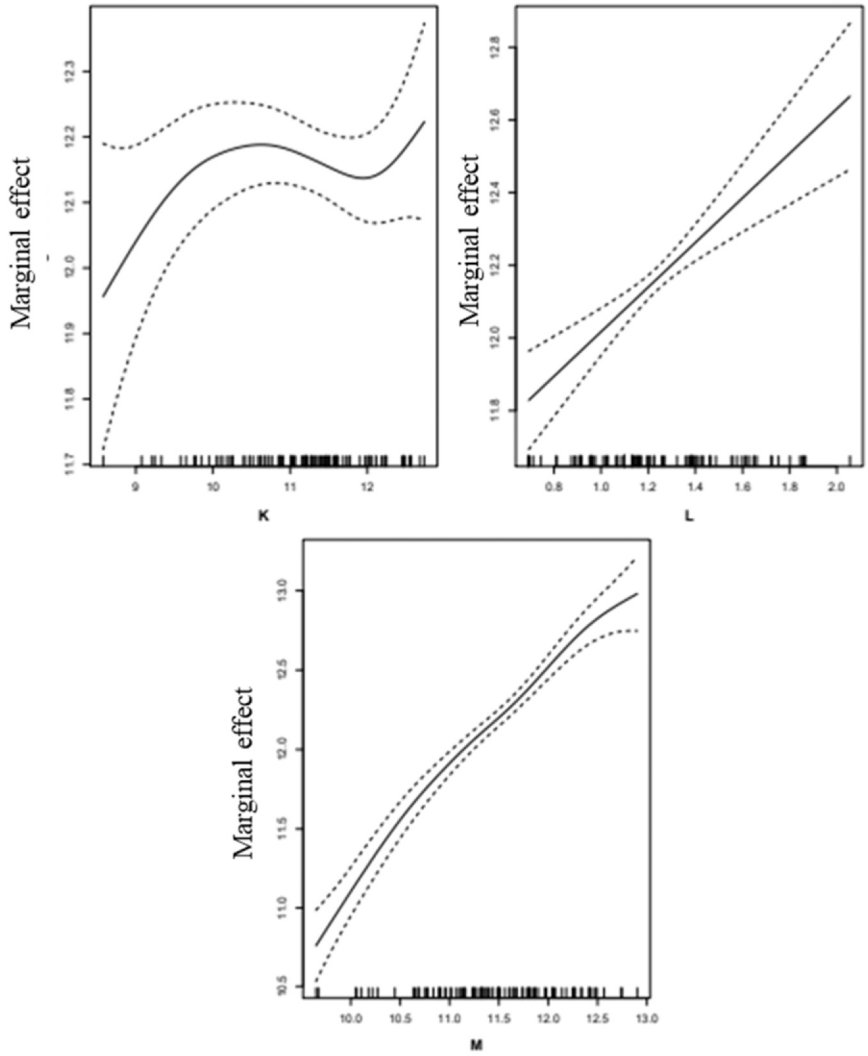

Fig. 2. Marginal effects with confidence intervals (dotted lines) of the estimated GAM frontier: the retail sector.

46%, respectively). This highlights that most of the efficient firms in the retail sample declare a turnover coherent with presumptions.

This analysis provides a useful tool to support the audit activities of fiscal authorities, since it could help to distinguish and rank the non-congruous firms under the threshold. For example, although firm **A** in Fig. 4 is highly efficient, the reported turnover value is significantly below the *BSS* congruity thresholds. This evidence could suggest to authorities that this firm is engaging in anomalous behaviour. In contrast, firm **B** has a non-congruity status that could in part be explained by the presence of a low efficiency score rather than non-compliant behaviour.[18]

Therefore, the combined analysis allows authorities to identify a ranking among firms and show which non-congruous taxpayers might need to be prioritised for additional controls.

To test the robustness of our findings, we replicated the combined analysis conditioned by quartiles using a single index.[19] Fig. 5 shows that smaller firms (i.e., the firms in quartiles I and II) have higher efficiency scores and are more likely to record turnovers higher than the estimated threshold. On the contrary, larger firms (quartiles III and IV) have higher spreads in terms of their

[18] In this case we fix the threshold efficiency value to the first quartile of the corresponding distribution.

[19] We divided the firms in the sample into quartiles based on the value of the sum of the standardised values of K and M rather than turnover (which was already used both to estimate the efficiency score and to compute the relative distance from the *BSS* congruity threshold). This choice originates from considering that the single index measures the level of inputs, and is therefore strongly correlated with revenues and could be used as a proxy for firm size.

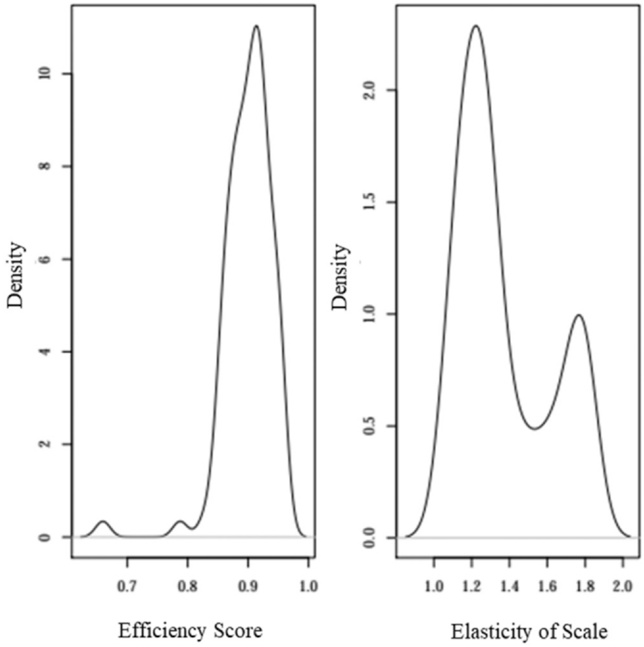

Fig. 3. Efficiency and elasticity of scale for the retail sector.

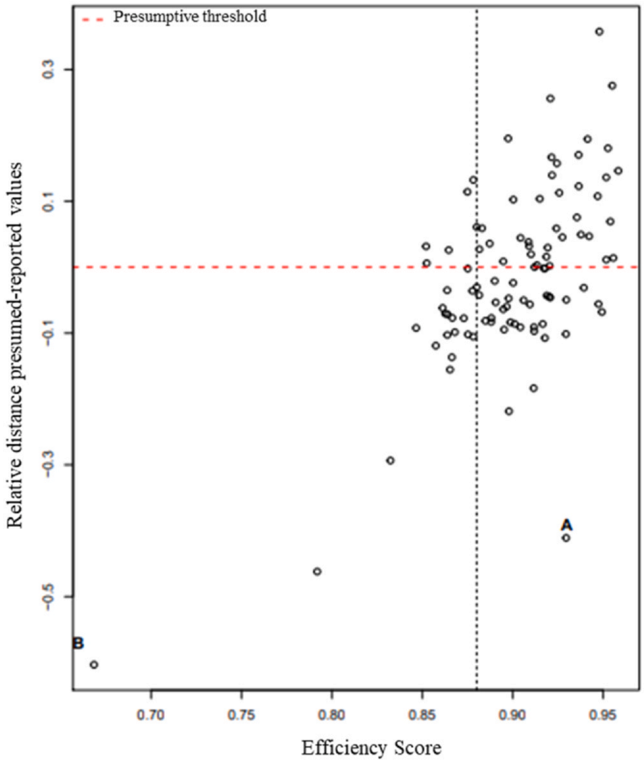

Fig. 4. Efficiency-congruity: the retail sector.

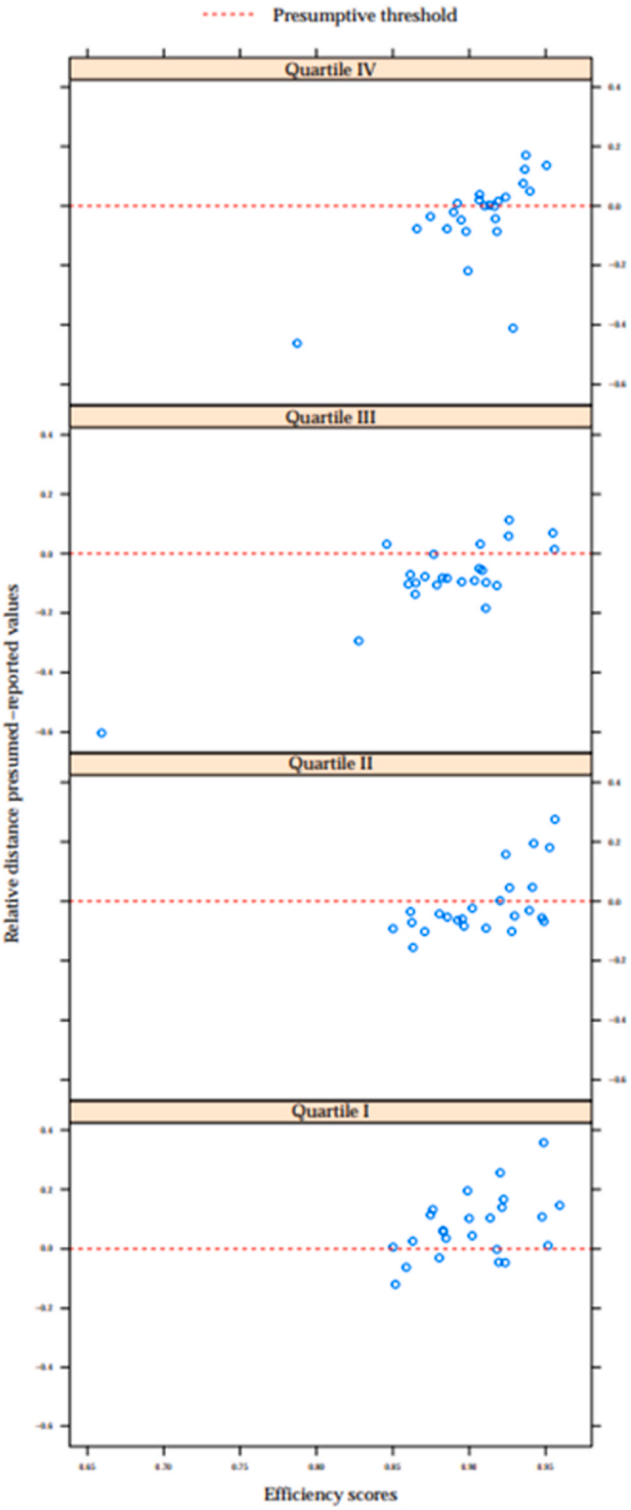

Fig. 5. Efficiency-congruity by size: the retail sector.

Table 2
Tax compliance in the retail sector.

	Turnover (€ x 1000)	*SSR*	*NW-SSR*
Full sample	230	1.836	0.952
I quartile	109	0.398	0.020
II quartile	181	0.274	0.088
III quartile	218	0.626	0.447
IV quartile	433	0.537	0.397

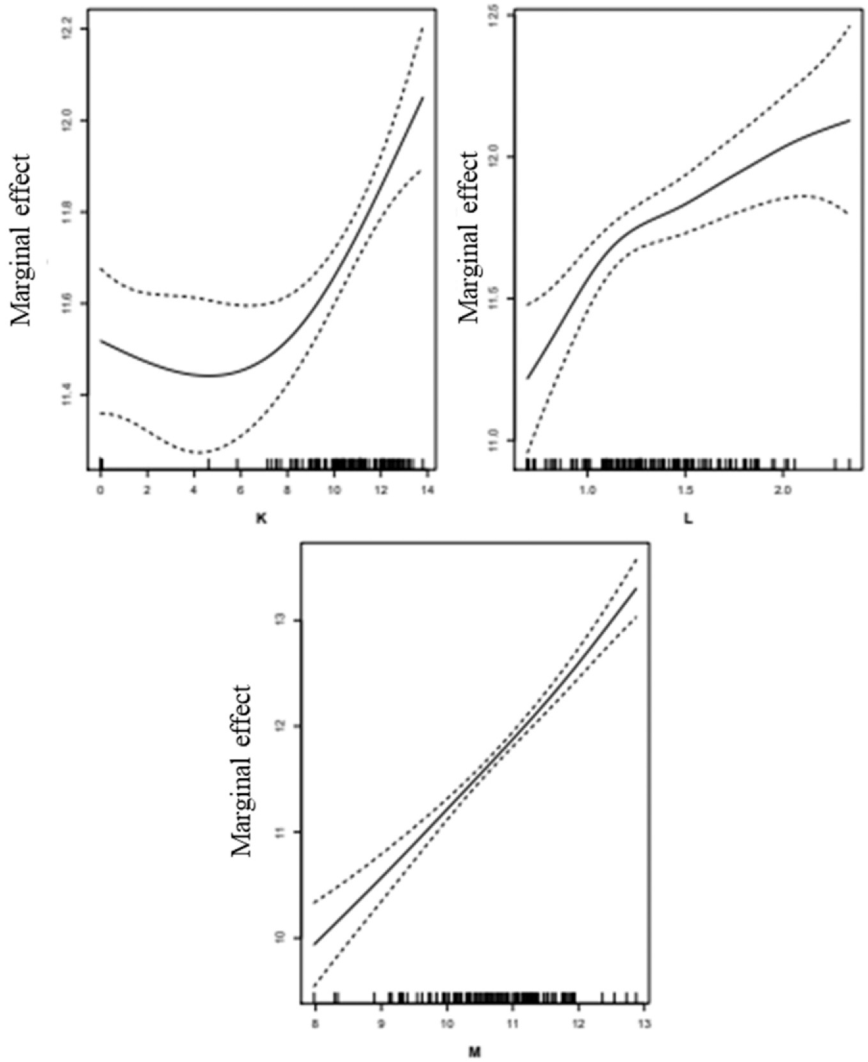

Fig. 6. Marginal effects with confidence intervals (dotted lines) of the estimated GAM frontier: the services sector.

efficiency scores and are more likely to declare turnover lower than the *BSS* estimated threshold. In particular, in quartile IV, several cases with under-reported turnover are not connected with low efficiency scores.

Finally, we provide some insights in terms of tax compliance by measuring the non-compliance level for the retail sample, both overall and by quartiles. The findings are summarised in Table 2.[20] The overall gap between the declared and presumed turnover estimated for the retail is equal to 1.836. If we compute the non-compliance measure by restricting the firms to the subset of taxpayers reporting a turnover

[20] Through sensitivity analysis, as described in Section 2.1, we verified that the main results do not depend on the specific value considered for inputs potentially affected by under-reporting. The results are reported in the Appendix.

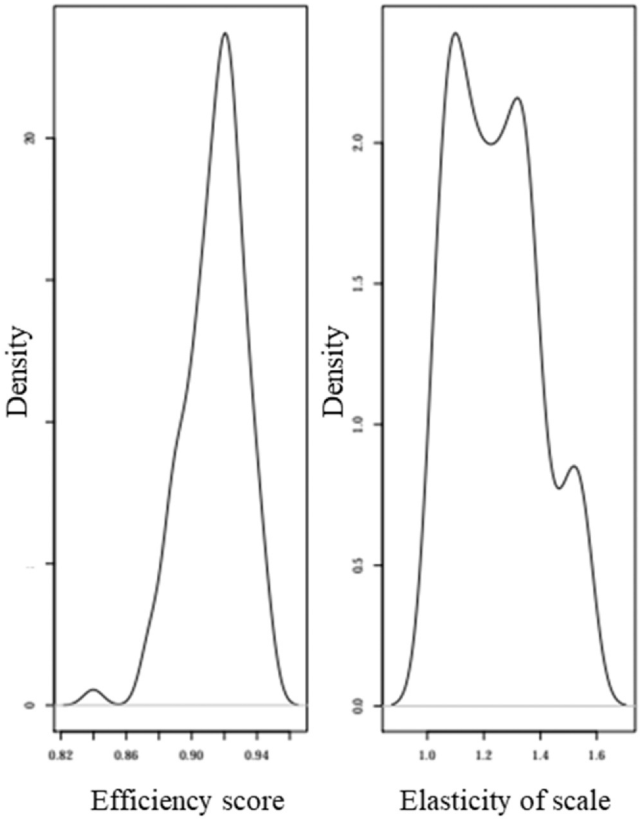

Fig. 7. Efficiency and elasticity of scale for the services sector.

lower than the presumed value, we find that the *NW-SSR* is almost half the tSSR. However, the most important result emerges when analysing the non-compliance measures by quartiles. The values of the *NW-SSR* for smaller firms (quartiles I and II) are much lower than those of larger ones (III and IV). This evidence, in line with the findings in Fig. 5, allows us to conclude that in the retail sample smaller firms are more compliant, while the high non-compliance measure for larger firms signals more frequent anomalous behaviours.

3.3. The services sector

As for the retail sample, the empirical application starts with the estimation of the labour variable, which is the input potentially affected by under-reporting. In particular, we specify the labour equation as follows:

$$L_i = \gamma_1 K_i + \gamma_2 M_i + \gamma_3 Gas_i + \nu_i i = 1, \ \dots, n. \tag{9}$$

We define the number of employees (L_i) as a function of the capital stock (K_i), the costs for materials (M_i) and the amount of gas used (Gas_i).[21]

In Fig. 6 we report the estimated frontier in terms of marginal effects, providing, as for the retail sector, evidence about the non-linearity of the production function estimated via $\psi(\cdot)$.

Fig. 7 shows that the efficiency scores range from 0.85 to 0.95 and that a high share of firms in this sample are quite efficient, since most of them report an efficiency score higher than 0.9 (approximately 86%). As for the retail sector, we report the distribution of the elasticity of scale, highlighting increasing return to scale for all firms.

Fig. 8 combines the efficiency scores estimated with the outcome of the *BSS* analysis in terms of the congruity for each firm.[22]

When we compare the reported turnover with the *BSS* threshold, both positive and negative differences emerge. However, if we focus on the tails of the estimated efficiency scores, we find interesting results. On the one hand, less efficient firms (i.e., those having efficiency scores lower than the first quartile) declare a turnover lower than the presumed *BSS* values. On the other hand, most of the more efficient firms declare a turnover higher than the *BSS* congruity threshold (86% of such firms are congruous). This suggests that in the services sample the *BSS* and SIF approaches lead to similar results: most of the non-congruous firms exhibit low efficiency scores, while most of the high efficiency firms are congruous based on the *BSS* presumptions.

[21] As for the retail sample, all these variables are expressed in logs.

[22] The results are reported in the Appendix.

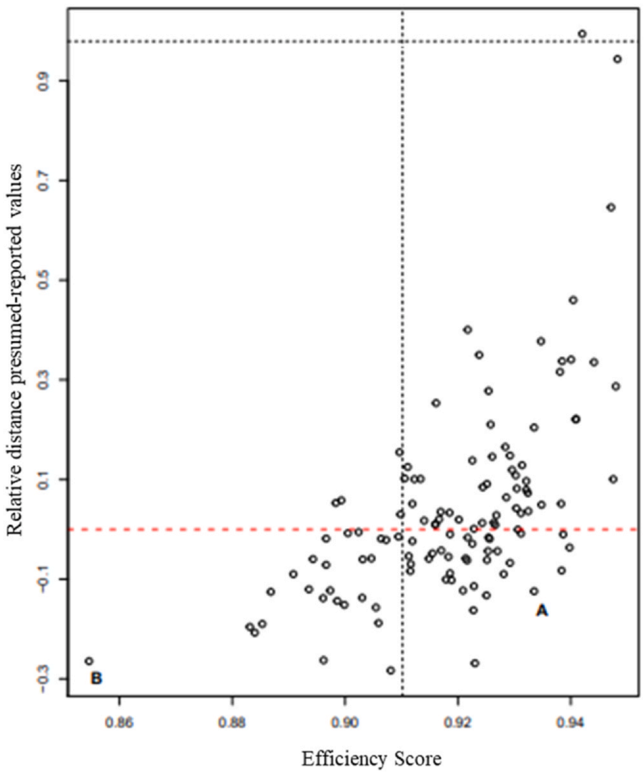

Fig. 8. Efficiency-congruity: the services sector.

However, as in retail, in the services sample, it is possible to find some examples of firms with diverging results. For example, despite firm **A** having a high efficiency score, it is non-congruous, thus highlighting the potential need for further investigation by the fiscal authority. In contrast, firm **B**, with a lower efficiency score and a high positive distance from the *BSS* presumed threshold, might not be among the firms prioritised for further investigation for potentially under-reporting data.

In line with the retail sector, we have replicated the analysis by quartiles using a single index, which is used as a proxy for firm size. The results are summarised in Fig. 9, which shows that the smallest firms (quartile I) exhibit high heterogeneity in terms of both measures: being above/below the presumed *BSS* threshold, but this is not necessarily linked to being ranked as having low/high efficiency. Larger firms belonging to quartiles II, III and IV are rather condensed in terms of both their efficiency performance and congruity results.

Finally, in Table 3, the analysis in terms of tax compliance shows that the *SSR* estimated for the whole services sample is 2.569.[23] The higher value of the non-compliance measure is concentrated in the first quartile, while lower values are recorded for the larger firms. Weighting the *SSR* for the efficiency scores, i.e. focusing on the *NW-SSR* measure, a split in non-compliance values related to firms' dimension emerges: the smallest firms (quartiles I and II) record the highest values, while the larger firms (quartiles III and IV) record the lower ones. These findings, in line with Fig. 9, highlight that in the services sector, smaller firms are less compliant than larger ones.[24]

3.4. A sector comparison

The empirical application of our integrated methodology to the two different sectors leads to quite different preliminary conclusions. By comparing Figs. 4 and 8 in terms of efficiency, we find that firms in the services sector are more efficient than those in the retail sector.

The combined efficiency-congruity analysis at the micro level shows that in both sectors negative differences between the reported and presumed turnover are quite frequent. However, the retail and services sectors differ in terms of their behaviour, with anomalous reporting being less ascribable to low efficiency in the retail than in the services sector.

Further, the sector tax compliance analysis shows that the firms in the retail sample exhibit lower tax compliance compared to the service sector. The *NW-SSR* computed for the services sector is lower than that computed for retail. The results are confirmed when the measure is weighted taking into account the overall number of firms in the corresponding sample ($0.952/102 \approxeq 0.009$ for the retail and $0.665/120 \approxeq 0.005$ for the service sector, respectively).

[23] We excluded the three largest positive distances considered as potential outliers.

[24] As in the retail sector, the sensitivity analysis showed that the main results do not depend on the specific value considered for inputs potentially affected by under-reporting. The results are reported in the Appendix.

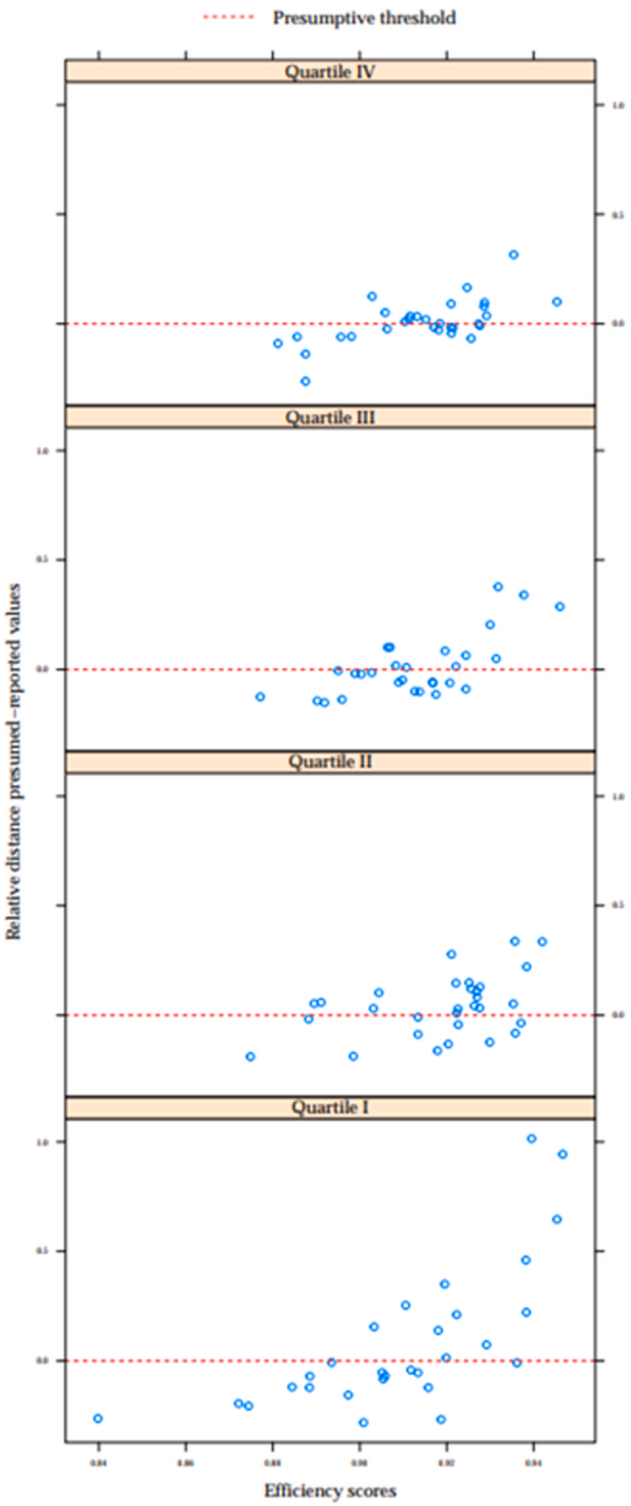

Fig. 9. Efficiency-congruity by size: the services sector.

Therefore, the combined efficiency-congruity microanalysis and the sector tax compliance considerations highlight the presence of more anomalous behaviours in the retail than in the services sample.

Our considerations have to be regarded as preliminary results due to the very small dimension of our samples and therefore require more investigation in order to extend the validity of the present findings.

Table 3
Tax compliance in the services sector.

	Turnover (€ x1000)	*SSR*	*NW-SSR*
Full sample	181	2.569	0.665
I quartile	58	0.938	0.325
II quartile	108	0.617	0.128
III quartile	166	0.555	0.116
IV quartile	381	0.300	0.095

4. Concluding remarks

This paper developed an integrated approach combining audit rules, presumptive taxation and a measure of technical efficiency estimated through the stochastic frontier analysis. This integrated methodology allows us to distinguish between firms' under-reporting that is potentially linked to taxpayer's inefficiency from a situation where discrepancies between reported and presumed values are more presumably ascribable to taxpayer's non-compliant behaviour. Further, we provide some considerations in terms of tax compliance. In particular, we computed a non-compliance measure based on the difference between reported and presumed revenues, supposing that the higher the value of under-reporting of revenues, the lower the tax compliance.

Our integrated methodology could provide a useful policy tool to support fiscal authorities' audit activities. Further, the empirical application of our integrated framework to different groups of taxpayers would allow a ranking in terms of higher/lower tax compliance, investigating if the level of tax compliance changes in relation to taxpayers' characteristics, such as, for example, different economic sectors.

For illustrative purposes we developed an empirical application based on the Italian *Business Sector Studies*. Given the very scant number of taxpayers analysed, it would be very interesting to test the robustness of this integrated methodology using a larger sample and moving from cross-section to panel analysis. The possibility to take into account individual-specific heterogeneity would allow us to study the dynamics and investigate taxpayers' behaviours more completely.

Disclaimer

The views expressed in the article are those of the authors and do not involve the responsibility of the corresponding institutions.

Acknowledgements

We would like to thank Hailun Way, Egidius Kundelis, and the participants at the 7th International Conference on The Shadow Economy, Tax Behaviour and Institutions, Brunel University (London, September 23–25 2021), for helpful comments. We thank the editor, the guest editors, and the two anonymous referees for many constructive comments and insightful suggestions that have helped us to improve the quality of this paper significantly. We also wish to thank Vincenzo Atella for helpful suggestions.

Appendix A

See Tables A1–A4.

Table A1
Labour regression estimation results.

(a) Retail Sector	
Costs for materials	0.333***
	(0.042)
Capital stock	0.071**
	(0.028)
Adjusted R^2	0.563
F Statistic	65.944***
(b) Services Sector	
Costs for materials	0.259***
	(0.037)
Capital stock	0.020**
	(0.010)
Gas consumption	0.018*
	(0.009)
Adjusted R^2	0.424
F Statistic	29.902***

Note: Significance codes: 0 "***", 0.001 "**", 0.01 "*".

Table A2
Gam frontier estimation results.

(a) Retail Sector	
Variables	**F**
Capital stock	1.237
Labour	25.639***
Costs for materials	41.957***
(b) Services Sector	
Variables	**F**
Capital stock	11.395***
Labour	4.639**
Costs for materials	63.283***

Note: Significance codes: 0 "***", 0.001 "**", 0.01 "*".

Table A3
Retail sector: sensitivity analysis of *NW-SSR*.

	Reference level $E(\hat{x}_k)$	$E(\hat{x}_k) - \hat{\sigma}$	$E(\hat{x}_k) - 2 \times \hat{\sigma}$
Full sample	0.952	1.004	1.052
I quartile	0.020	0.021	0.021
II quartile	0.088	0.090	0.091
III quartile	0.447	0.485	0.520
IV quartile	0.397	0.408	0.420

Table A4
Service sector: sensitivity analysis of *NW-SSR*.

	Reference level $E(\hat{x}_k)$	$E(\hat{x}_k) - \hat{\sigma}$	$E(\hat{x}_k) - 2 \times \hat{\sigma}$
Full sample	0.665	0.706	0.800
I quartile	0.325	0.349	0.398
II quartile	0.128	0.133	0.149
III quartile	0.116	0.122	0.138
IV quartile	0.095	0.102	0.116

References

Aigner, D., Lovell, C.A.K., Schmidt, P., 1977. Formulation and estimation of stochastic frontier production function models. J. Econ. 6, 21–37.
Allinhgam, M.G., Sadmo, A., 1972. Income tax evasion: a theoretical analysis. J. Public Econ. 1, 323–338.
Alm, J., 2019. What motivates tax compliance? J. Econ. Surv. 33 (2), 353–388.
Alm, J., Cronshaw, M.B., McKee, M., 1993. Tax compliance with endogenous audit selection-rules. Kyklos 46 (1), 27–45.
Alvarez, R., Crespi, G., 2003. Determinants of technical efficiency in small firms. Small Bus. Econ. 20 (3), 233–244.
Arachi, G., Santoro, A., 2007. Tax enforcement for SMEs: lessons from the Italian experience? eJ. Tax Res. 5, 224–242.
Berger, A.N., Bonaccorsi di Patti, E., 2006. Capital structure and firm performance: a new approach to testing agency theory and an application to the banking industry. J. Bank. Financ. 30, 1065–1102.
Bottasso, A., Sembenelli, A., 2004. Does ownership affect firms' efficiency? Panel data evidence on Italy. Empir. Econ. 29 (4), 769–786.
Bucci, V., 2020. Presumptive taxation methods: a review of the empirical literature. J. Econ. Surv. 32 (2), 372–397.
Dube, G., 2018. The design and implementation of minibus presumptive taxes. Serv. Ind. J. 38, 723–741.
Fan, Y., Li, Q., Weersink, A., 1996. Semiparametric estimation of stochastic production frontier models. J. Bus. Econ. Stat. 14, 460–468.
Ferrara, G., 2020. Stochastic frontier models using R. Handbook of Statistics – Financial, Macro and Micro Econometrics Using R. 42. pp. 299–326.
Ferrara, G., 2011. Statistical Models for Tax Monitoring: Open Problems and Methodological Advances. Book of short papers of Ph.D. thesis in Statistics and Applications. Italian Statistical Society, p. 2.
Fiorio, C., Iacus, S., Santoro, A., 2013. Taxpaying Response of Small Firms to an Increased Probability of Audit: Some Evidence from Italy. Working Papers 251, Department of Economics, University of Milano-Bicocca.
Gokalp, O.N., Lee, S.H., Peng, M.W., 2017. Competition and corporate taxation: an institution-based view. J. World Bus. 52 (2), 258–269.
Hastie, T.J., Tibshirani, R.J., 1990. Generalized Additive Models. Chapman & Hall, London, pp. 1990.
Heshmati, A., 2003. Productivity growth, efficiency and outsourcing in manufacturing and service industries. J. Econ. Surv. 17 (1), 79–112.
Kasper, M., Alm, J., 2022. Audits, audit effectiveness, and post-audit tax compliance. J. Econ. Behav. Organ. 195, 87–102.
Kirchler, E., 2007. The Economic Psychology of Tax Behaviour. Cambridge University Press, Cambridge.
Kumbhakar, S., Park, B., Simar, l., Tsionas, E., 2007. Nonparametric stochastic frontiers: a local maximum likelihood approach. J. Econ. 137, 1–27.
Kumbhakar, S.C., Lovell, C.A.K., 2000. Stochastic Frontier Analysis. Cambridge University Press, Cambridge.

Logue, K., Vettori, G., 2011. Narrowing the tax gap through presumptive taxation. Columbia J. Tax Law 2, 101–146.
Margaritis, D., Psillaki, M., 2007. Capital structure and firm efficiency. J. Bus. Financ. Account. 34 (9–10), 1447–1469.
Martins, A., Sa, C., 2018. The computation of taxable income when accounting numbers are not reliable: a note on presumptions. Int. J. Law Manag. 60 (2), 543–562.
Meeusen, W., van den Broeck, J., 1977. Efficiency estimation from cobb-douglas production functions with composed error. Int. Econ. Rev. 18, 435–444.
Memon, N., 2013. Looking at Pakistani presumptive income tax through principles of a good tax? eJ. Tax Res. 11 (1), 40–78.
Pisani, S., 2004. Il triathlon degli studi di settore. Agenzia delle Entrate. Roma.
R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Reinganum, J.F., Wilde, L., 1985. Income-tax compliance in a principal agent framework. J. Public Econ. 26 (1), 1–18.
Santoro, A., Fiorio, C., 2011. Taxpayer behavior when audit rules are known: evidence from Italy. Public Financ. Rev. 39, 103–123.
Santoro, A., 2008. Taxpayers' choices under studi di settore: what do we know and how can we explain it? Giornale degli Economisti e Annali di Economia, 67, pp. 161–184.
Sickles, R., Zelenyuk, V., 2019. Measurement of Productivity and Efficiency: Theory and Practice. Cambridge University Press, Cambridge.
Slemrod, J., 2007. Cheating ourselves: the economics of tax evasion. J. Econ. Perspect. 21 (1), 25–48.
Slemrod, J., 2019. Tax compliance and enforcement. J. Econ. Lit. 57 (4), 904–954.
Slemrod, J., Weber, C., 2012. Evidence of the invisible: toward a credibility revolution in the empirical analysis of tax evasion and the informal economy. Int. Tax Public Financ. 19 (1), 25–53.
Slemrod, J., Yitzhaki, S., 1994. Analyzing the standard deduction as a presumptive tax. Int. Tax Public Financ. 1 (1), 25–34.
Slemrod, J., Yitzhaki, S., 2002. Tax avoidance, evasion and administration. In: Auerbach, A.J., Feldstein, M. (Eds.), Handbook of Public Economics III. Elsevier, Amsterdam.
Tedds, L.M., 2010. Keeping it off the books: an empirical investigation of firms that engage in tax evasion. Appl. Econ. 42 (19), 2459–2473.
Torgler, B., 2016. Tax compliance and data: what is available and what is needed. Aust. Econ. Rev. 49 (3), 352–364.
Vidoli, F., Ferrara, G., 2015. Analyzing italian citrus sector by semi-nonparametric frontier efficiency models. Empir. Econ. 49, 641–658.
Wood, S.N., 2003. Thin plate regression splines. J. R. Stat. Soc. B 65 (1), 95–114.
Wood, S.N., 2006. An Introduction to Generalized Additive Models with R. CRC Press, Boca Raton, FL.
Yitzhaki, S., 2007. Cost-benefit analysis of presumptive taxation. FinanzArchiv/Public Financ. Anal. 63 (3), 311–326.