Tax Gap and Redistributive Aspects Across Italy

Alessandra Braiotta (Agenzia delle Entrate, Italy)
Alfonso Carfora (Agenzia delle Entrate, Italy)
Rosaria Vega Pansini (Agenzia delle Entrate, Italy)
Stefano Pisani (Agenzia delle Entrate, Italy)

Paper Prepared for the IARIW 33rd General Conference
Rotterdam, the Netherlands, August 24-30, 2014

Session 7A
Time: Friday, August 29, Morning
Resumè

Abstract ......................................................................................................................................................... 1
Introduction ................................................................................................................................................... 2
An overview of theoretical approach ............................................................................................................. 3
Methodological issues in tax gap measurement ............................................................................................ 4
The decomposition of tax gap by gross profit and wages ............................................................................. 5
Bottom up approach to decompose the ‘top-down’ tax base ........................................................................ 7
Propensity score matching and tax evasion estimation .................................................................................... 8
Data ............................................................................................................................................................... 9
Main results ................................................................................................................................................. 10
Decomposition of IRAP BIND ...................................................................................................................... 10
Estimation of Tax evasion of self-employees through propensity score analysis ....................................... 12
Concluding remarks ..................................................................................................................................... 17
References ................................................................................................................................................... 19

Abstract

The aim of the paper is twofold: first, we contribute to the debate on the estimation of tax gap proposing a novel methodology to empirically assess and disaggregate the (total) tax gap; second, we analyze the effects of tax evasion in terms of primary distribution of income.

The novelty of the paper relates to the methodology used to estimate tax gap based on an original integration of the top-down and the bottom-up approach. In fact, top-down approach, based on the comparison and cross-checking (matching) between national accounts and administrative fiscal data, is employed to obtain the estimation of the total gap in the tax base (Braio tta et al., 2013). The bottom-up approach, based on results from fiscal audits, is used to derive single components of the tax gap related to single taxes and different taxpayers (Feinstein, 1999; Danish Tax and Customs Administration, 2006; IRS, 2012).

Furthermore, the use of micro data from fiscal audits allows to examine how tax evasion modifies the progressivity of the tax system as well as vertical and horizontal equity in the distribution of income. The analysis is conducted for Italy at the regional level in order to analyze the territorial distribution of total tax gap.
Introduction

Previous and recent literature on tax avoidance has always clearly recognized the negative effects of tax evasion on economic development, fair competition, public governance and social cohesion (among others, Roubini N., Sala-i-Martin, X., 1992; Caballe and Panades, 2000). In the recent economic downturn, more and more attention has been devoted by national and international policy makers to instruments for reducing the magnitude of evasion and the way in which benefits from increasing tax compliance can be used to enhance socio-economic development. These issues acquire a certain relevance in the context of EU member States, and in particular, in Italy.

In the light of the current economic and financial crisis, it has become urgent that EU countries rethink the way in which public services are financed. Moreover, giving the Fiscal Compact treaty, countries with a high level of public debt, like for example Italy, have stricter conditions to finance public spending and investments. Finally, considering that policy makers of EU member States cannot fully use monetary policy instruments to influence GDP, fiscal policy and the administration of public revenues is the (main) countercyclical tool. In such political and economic context, it is crucial to preserve the efficiency of tax collection as it represents the main way to ensure the sustainability of the State budget. Thus, as enhancing tax compliance directly improves a country’s ability to collect resources for public goods provision and income redistribution, a more comprehensive approach to estimate the size of resources subtracted from public revenues due to tax evasion is necessary. In addition, considering the size of tax gap within the assessment of state budget is an essential condition in order to harmonize the welfare and taxation systems among countries with different characteristics and to avoid any form of arbitrage or unfair competition. (European Commission, 2011; OECD, 2002, 2008). Lastly, it is also important provide disaggregated figures for the tax gap, in order to show the contribution of the evasion of each tax to the total evasion in order to analyses the overall coherence of fiscal system

Beside reducing the space for fiscal policy, tax evasion can modify the distributive impact of tax legislation, altering the mechanism through which income is produced and distributed to factors of production and then to taxpayers. In fact, the possibility to evade can create distortions in the primary distribution of income, i.e. across profits and labor income, and in the secondary distribution of income, between different types of individuals, i.e. across employees and self-employed taxpayers. Thus, tax evasion has impact on horizontal equity, as people with the same level of income face different tax burdens, and vertical equity, as taxpayers with higher capacity to pay can have also higher possibility to avoid tax payment (CEPAL, 2012).

The aim of the paper is twofold: first, we contribute to the debate on the estimation of tax gap proposing a novel methodology to empirically assess and disaggregate the (total) tax gap into components related to single taxes; second, we analyze the effects of tax evasion on primary distribution of income and on different types of taxpayers. Considering a high degree of black market labor in the Italian economy, we intend to provide empirical evidence about the very great influence exerted on public finance flow, and in particular on the distribution of benefits arising from the evasion of income earners as employee or self-employed.

Our analysis aims at answering two main questions. First, is it possible to estimate the total tax gap and then derive disaggregated estimates for single taxes and for single types of taxpayers? As for different taxpayers, is there a difference in their compliance behavior that influence the magnitude of tax gap? Second, in the presence of tax evasion, how the primary distribution of value added between profits and wages changes?

We will try to answer to the previous questions using a novel approach in the estimation of tax gap that originally integrates top-down and bottom-up methods to assess the size of the phenomenon. Top-down approach allows evaluating the total tax gap connected to the production of goods and services, while using bottom-up it is possible to distinguish between tax gap related to individual firms from that of companies. Moreover, bottom-up approach allows also distinguishing tax gap associated with different types of

---

1 The OECD defines tax evasion as ‘illegal arrangements where liability to tax is hidden or ignored’, while tax avoidance is described as ‘an arrangement of a taxpayer’s affairs that is intended to reduce his liability and that although the arrangement could be strictly legal it is usually in contradiction with the intent of the law it purports to follow’. (OECD, 2002)
individual firms as compatible with the Italian fiscal law, thus among self-employees and small individual companies. Empirical computations rely on different sources of data: (a) national accounts, (b) administrative data on individual and companies’ income tax returns and (c) results from fiscal audits conducted by the Italian Revenue Agency. The analysis is conducted for Italy at the regional level in order to analyze also the territorial distribution of total tax gap and its components.

The paper is organized as follows. Section 2 recalls the main literature on the estimation and disaggregation of total tax gap. Section 3 outlines the main methodological issues related to the assessment of tax evasion and sketches the theoretical framework adopted in the paper. Section 4 shows some preliminary results applied to Italian regions and different taxpayers. Finally, Section 5 outlines some conclusions and policy implications.

**An overview of theoretical approach**

Literature on tax evasion and tax compliance is extensive and comprehensive, mainly concerned with the definition and estimation of the size of shadow economy (see, among others, Schneider - Enste, 2000). In the majority of cases, empirical studies try to assess the size of gap related to single taxes or to the use of irregular job (Reckon, 2009; Smith – Keen, 2007; Schneider, 2012, 2013; D’Agosto et al., 2013). Few studies try to estimate the total size of tax gap and then disaggregate it into its components related to the different taxes. Thus, in this cases, it is not easy to compare and derive conclusions on the relative contribution of evasion of single taxes on total tax gap or to study the evading behavior of different taxpayers, (e.g. firms versus employees).

With reference to the Italian case Braiotta et al. (2013) provides a decomposition of the tax base into the components of gross profits and the cost of labor, analyzing the effects of tax evasion on the primary distribution of income. Our paper innovates both in further disaggregating the tax gap and integrating top down and bottom up approaches for the estimation of total tax gap.

Due to difficulties related to the estimation of gap for employees, more attention has been devoted to the gap in profits or in the total of value added produced by firms. In the attempt to concentrate on tax compliance of firms, literature has mainly focused on: the relationship between specific forms of market and different propensities to compliance (Marrelli, Martina, 1998); corporate governance (Chen, Chiu, 2002; Santoro, 2005); the dynamics of non compliance connected to the business dimension adopted (Rice, 1992; Farina, Pisani, 2009). For example, small and medium entrepreneurs operate similarly to self-declaring taxpayers, however firms don't seem to be influenced by social effects like guilt or shame if compared to individuals; other factors, such as the perceived usefulness and quality of goods and services provided by the government, affect tax toleration and improve tax compliance (Hibbs and Piculescu, 2010).

Among the components of total tax gap, beside profits non registered workers has a crucial role, because tax evasion derived from black market labor can itself be distinguished into remuneration of irregular employees and gross profits derived from the use of irregular job into the production process. As shadow economy and tax evasion, also non registered workers has been extensively analyzed both in national and international context (ISTAT, 2010, 2011, OECD 2002 Giovannini, 2010) and under different perspectives.

In particular, several theoretical and empirical contributions have focused on the relationship between the structural characteristics of the labor market (e.g. rigidity and minimum wages), on the different measures of fiscal policy (level of direct taxation and the tax wedge), and on the supply of labor in the informal sector (Basu et al, 2011, Lemieux et al. 1994, Frederiksen et al. 2005).

---

2 See, among others, Andreoni et al, 1998; Slemrod – Yitzhaki, 2002; Slemrod, 2007. Besides methodological and empirical papers, there is a wide literature that reviews the principal methodologies and results on tax gap assessment. See, among the more recent, Gemmell and Hasseldine (2012) and Alm (2012).
Furthermore, theoretical literature on tax evasion has been concentrated on under-reporting of self-employment earnings (Hurst et al., 2010); conversely, the typical models assume no evasion on wage and salary income (Feldman and Slemrod, 2007). Whereas the withholding from wages and salaries earned exists it makes more difficult under reporting (Slemrod and Weber, 2011). However, this literature has neglected the presence of a potential agreement between employers and employees who cheat on tax authorities and keep a fraction of earnings undeclared (Sandmo, 2011). This collusion encourages different forms of tax evasion: between off-the-books versus registered work; between full non-compliance and full compliance, and “on-the-books” salaried employment, who hide a part of payments.

An additional element of originality of this work is to integrate bottom up and top down approach to disaggregate total tax gap. Still, bottom up techniques are always more used to assess the total tax gap of specific types of taxpayers. In fact, the availability of data at the micro level from fiscal audit allows to fully exploit information on the characteristics of audited taxpayers in order to derive conclusions on their fiscal behavior and on the size of tax gap connected to different phase of the production process or different type of income earned. Empirical models based on bottom-up techniques dealing with US tax audits, report a rate of under-reporting of income from dependent employment (0.5%) much lower than for self-employment income (58.6%) under Taxpayer Compliance Measurement Program (TCMP), (Slemrod and Yitzhaki, 2002) and an estimated 57% of self-employment income underreported, compared to 1% of wages and salaries under the National Research Program (NCP) (Slemrod, 2007). For the Danish firms the marginal tax rate has at most a small positive substitution effect on tax evasion for individuals with substantial self-reported income (Kleven et al, 2011). In Italy, it has been estimated a rate of under-reporting of self-employment income around the median of the distribution at 27.7%, compared to 1.9% for income from wages and salaries (Fiorio and D’Amuri, 2005).

Studies at the micro level have been also employed to analyze how the distribution of income changes due to the presence of tax evasion. Using data from the TCMP, Bishop et al. (2000) found that while the effects of tax evasion on vertical equity are small, there is a substantial effect on horizontal equity generated by non compliance. More recently, Freire-Seren and Panades (2008) analyze the distortive effect of tax evasion on the redistribution due to tax progressivity. Using the Lorenz curve approach and data on Spanish Personal Income Tax, they compare the distribution of after-tax income without evasion, the distribution of expected income with evasion and the distribution of after-tax declared income. They find that tax evasion generates distortions since the distribution of after-tax income without evasion dominates in the Lorenz sense that of after-tax declared income before auditing.

The same distortive effects is also found by Florio and D’Amuri (2006). They investigate the effects that different probabilities of evading taxation have on the distribution of income of different taxpayers across different income levels and types of income. Increasing possibility to evade might cause distortions in the distribution of income and effects on poverty, income inequality and tax revenues. It can be the case that employees can decide to move to self-employed position and lower their tax burden or people employed in the hidden market can be discouraged to enter in the official labor market.

Bloomquist (2003) also investigates on the link between tax evasion, detection probability and income inequality. Using US data for the period 1947-2000, he finds that higher income inequality lower detection probability changing the composition of income from labor to investment income. Moreover, income inequality increases compliance opportunity costs if relatively wealthier taxpayers perceive an unbalance between their tax burden and public sector benefit. The author then concludes that there is a positive correlation between income inequality and tax evasion.

Methodological issues in tax gap measurement

The aim of the paper is to disaggregate total tax gap into gap related to single types of taxes using an integrated top-down and bottom up estimation approach. This disaggregation is functional to show how the
incidence of tax evasion on different types of income alters the primary income distribution. Moreover, bottom up approach will allow to further disaggregate tax gap for different types of taxpayers to show if tax evasion impacts also horizontal equity.

For the specific purpose of this paper, we focus our attention on the Tax on Regional Productive Activities (hereafter IRAP, as per the Italian acronym) ³.

The main features of this tax are:

- a large number of taxpayer spanning the operators comprised in the assessment of the GDP almost in full;
- a substantially flat rate (around 4%)
- a definition of tax base very similar to that of value added at factor cost, as stated by the system of national accounts⁴;

The similarity between the IRAP tax base and value added from national account is essential in the study of the tax gap. In fact, the latter is the basic unit that determines the GDP and, therefore, it contains all the types of income that generate the change of the wealth of a country. It follows that the IRAP tax base encompasses much of the tax base resulting from the production of goods and services.

In addition, the large number of taxpayers subjected to the tax means that the IRAP tax base gap represents a macro indicator of the value added eventually concealed from tax authorities. In this case, the tax gap is not so important to focus on as the gap of the tax base. The procedure involves the calculation of the evasion for both the tax base (base gap) and the corresponding tax (the tax gap).

The approach used in the estimation of the IRAP tax gap follows the ‘top-down’ approach, based on the comparison of the tax base inferred from the IRAP tax income return forms (declared income tax base, hereafter BID) with data from National Accounts on value added at factor cost⁵, which represents the potential tax base (hereafter BIT). To make such a comparison is necessary to harmonize the two quantities from the point of view of the definitions and classifications adopted and of the resulting discrepancy which is attributable solely to the IRAP base not declared (undeclared income tax base, hereafter BIND). Then, by applying the statutory tax rates it is possible to calculate the IRAP tax gap.

The decomposition of tax gap by gross profit and wages

As stated before, the design of the IRAP makes the base of this tax almost a proxy of the value added at factor costs. Moreover, once the necessary reconciliations have been made, it can be considered as the sum of labor costs, as the remuneration of employees (wages)⁶, and a residual part related to gross profits, as the remuneration of firms and small entrepreneurs. This decomposition allows to disentangle the components of the tax base, and then of the tax gap, related to employed labor (wages) and to self-employed activities. Being able to distinguish these two components can also help to shed light to the motivation to evade of different taxpayers. As observed in Farina and Pisani (2009), it is probable that self-employed taxpayers decide whether and how much to evade following a model of behavior based on the convenience to avoid taxes, for example, à la Allingham and Sadmo, i.e. dependent on economic benefits to evade and costs

---

³ IRAP is specific of Italian tax system and it has been introduced starting from 1997.
⁵ For more details, see Braiotta et al (2013).
⁶ According to National Accounts definition labor costs are represented by: a) gross wages and salaries including contributions paid by employees, direct remunerations, bonuses and allowances paid by an employer in cash to an employee in return for work done, payments to employees saving schemes, payments for days not worked and remunerations in reward with food, drink, fuel, company cars, etc. and (b) other labor costs including employers' social contributions plus employment taxes less subsidies received by employers.
associated to the probability of being audited; employees, especially low-skilled workers, do not always decide autonomously to evade, but conditionally on the environment in which they operate.

As regard to the definition of labor cost, firms correspond wages and salaries (inclusive of contributions paid by the employee) and employers’ social contributions to regular workers (Wd), while those irregular perceive only the gross wage and salary (Wnd). The distinction of labor costs into regular and irregular allows to decompose both the declared (BID) and the undeclared base (BIND).

Once IRAP tax base gap (BIND) has been obtained, we remove the undeclared work (Wnd) and get the hidden residual part (undeclared gross profits, GPnd). With this decomposition it is possible to analyze the functional distribution of income in presence of tax evasion and define the link between undeclared work and tax evasion. (Farina and Pisani, 2009). Figure 1 offers a schematic representation of the methodology employed to decomposed the declared and undeclared IRAP tax base.

**Figure 1: IRAP tax base and primary distribution of income**

![Diagram](image_url)

*Source: authors’ elaboration based on Braiotta et al. (2013)*

Figure 1 helps also to visualize that base (declared and undeclared) of direct taxes can also be derived from IRAP tax base. In fact, subtracting from BID and BIND IRAP the cost of labor, we obtain a measure for gross profits that is very close to the tax base of direct taxes, namely IRPEF (personal income tax) paid on profits of self-employees and small-size companies, and IRES (corporate income tax) paid on profits by medium-size and limited companies. As well as the tax base, this decomposition is made possible also for the tax gap, distinguishing between the component of avoided profits by small firms and that of relatively bigger-sized companies.

We want to test if the presence of tax evasion operates a redistribution of primary income in favor of the component related to the remuneration of capital (profits) instead of labor (wages). The redistribution is, in fact, based on the fact that tax base evasion causes a change in the incidence of wages and profits between BID and BIND, so that it results in a redistribution of value added quotas from workers to entrepreneurs.

Beside the effects on primary distribution of income, as noted before, tax evasion can have serious impacts also on secondary distribution between different types of taxpayers receiving the same type of remuneration (wages or profits). Thus we can implement further the decomposition of tax base with respect to the component of direct taxes paid by self-employees and small-size companies (IRPEF). As stated by the Italian fiscal law, IRPEF burdens on profits of self-employees (SE) and small size firms and companies (SC) (namely ‘persone fisiche’), while IRES on profits of the limited companies, both big structured companies (namely ‘società di capitali’) and medium-sized firms with a simpler structure of business (namely, ‘società
di persone’) (LTD). Given that taxpayers subjected to IRPEF and IRES can have a different propensity and probability to evade, it is also useful to decompose the base and the gap of these two taxes into components related to single types of taxpayers. Figure 2 sketches the method to derive a decomposition of the RPEF base from BIND IRAP. Here we focus on the disaggregation of IRPEF in order to show the component of evasion of profits from self-employees from the one from small firms. The purpose is to show if tax evasion operates a redistribution of income also among different types of taxpayers subjected to the same tax, as such to influence horizontal equity.

**Figure 2: Decomposition of total undeclared Value Added**

The methodology used to decompose the IRAP tax base is shown in figure 1 and it is based on the top down approach. One of the shortcomings of this approach is the use of data at the macro level that does not allow to derive a robust estimation of the components of tax gap. Estimation of the cost of labor, regular (Wd) and irregular (Wnd) is provided by the Italian National Statistics Institute (ISTAT). Moreover, using data on firms’ income tax returns, it is possible to derive the BID related to profits and disaggregate it into the IRPEF and the IRES components. At the present, the same disaggregation for the BIND has been possible only applying hypothesis on the magnitude of tax gap derived from information on declared income. As business decisions can be tremendously different from compliant and non compliant taxpayers, such an hypothesis can appear restrictive.

The original contribution of the paper states in the integration of results from top-down and-bottom up approaches to decompose the total tax base and gap. Precisely, using data from fiscal audits and a resulting measure of tax evasion intensity, it has been possible to divide the total IRAP BIND into BIND IRPEF and BIND IRES. Moreover, data at the micro level have also been employed to decompose the BIND IRPEF for self-employees (SE) and small firms (SC), (Figures 1 and 2).

Bottom-up approaches for estimating tax evasion have undoubtedly advantages linked to the richness of information on the characteristics of audited taxpayers and their business and the undeclared tax base (Farina and Pisani, 2009). Previous examples applied to Italy are the analyses of Bordignon and Zanardi (1997), Marino – Zizza, 2010, and Fiorio et al. (2013). These studies are referred only to a specific tax or type of individuals due to limited availability of data on fiscal audits.

---

1 For a detailed review of the estimation of regular and irregular cost of labor, see Braiotta et al. (2013) and ISTAT (2011).
2 We should also underline that the hypothesis to apply to non declared tax base some of the information available for the declared base is the only hypothesis possible if other reliable source of data are not available.
3 For an extensive review see Gemmell and Hasseldine (2012), Palmieri (2004) and Marino Zizza (2010)
Bottom-up approaches has also a major disadvantage represented by selection bias when audited taxpayers are not randomly selected, as in the case of controls performed by the Italian Revenue Agency. Thus, any assessment with the use of the (sole) data from fiscal audits will then results in a biased and under/overestimated tax evasion. In order to correct this bias, two different methodologies have been applied.

The first is a post-stratification of average per capita tax evasion to derive the relative size of IRPEF and IRES gap. The methodology to derive estimates of base and tax gap for both taxes requires different steps. First, as data on audited taxpayers are not a representative sample of the population, a suitable clustering has been applied based on information of the dimension of business and macro region. Second, a preliminary analysis of the distribution of amounts of taxes has been performed together with outliers control. Third, average per capita tax evasion per strata has been estimated and imputed in cases where information was missing. Fourth, a post stratification procedure has been applied to obtain estimates of tax evasion for the entire population using the inverse of the probability to be audited per stratum as weights. Finally, a measure of tax evasion intensity has been derived as the ratio between the BIND and the BID at the regional level.

The second bottom-up method employs the estimation of a propensity score matching model that will be presented in the next session.

**Propensity score matching and tax evasion estimation**

As part of bottom-up techniques to derive an unbiased estimation of the tax gap, matching methods have been extensively used. In our analysis, they have been applied to further decompose the evaded tax base for IRPEF into its components related to different taxpayers. Among different possible methods, propensity score matching (PSM) has been applied. It constructs a statistical comparison group based on a model of the probability of participating in the treatment (probit), using observed characteristics of both treated and not treated individuals. Participants are then matched to non-participants on the basis of propensity scores (Rosenbaum and Rubin, 1983). The propensity score is the conditional probability of assignment to the treatment given a vector of observed covariates and it is generated for each variable, with missing values to indicate the probability of that observation being missing.

When data have a monotone missing pattern, some empirical steps are necessary to impute values for each variable \(Y_i\) with missing values (Rubin, 1997):

1. Create a variable \(R_i\) with the value 0 for observations with missing \(Y_i\) and 1 otherwise.
2. Fit a probit regression model of type:

   \[
   x_i = \alpha + \sum_{j=1}^{n} \beta_j x_{ij} + \epsilon_i
   \]  

3. Create a propensity score for each observation to indicate the probability, \(p_i\), of being missing where:

   \[
   p_i = \Phi(\pi_i)
   \]  

and \(\Phi(\cdot)\) is the cumulative distribution function of a standard Normal distribution.

4. Assuming conditional independence and a sizable overlap in propensity scores between participants and matched nonparticipants, select a matching criteria to assign treated individuals to non-treated ones on the basis of the propensity scores.

Empirical analysis is carried out with the twofold aim of:

1. Correcting the potential selection bias in the auditing taxpayers’ selection criteria that might affect results of post-stratification of tax evasion;
2. Imputing the unbiased in-sample results of tax evasion estimation to the population of taxpayers to obtain a complete measure of the entire phenomenon.

The idea is to distinguish the selected sample of taxpayers in audited ($R_i=1$) and not audited ($R_i=0$) individuals, then to calculate for each individual a propensity score to match non-treated individuals to treated ones and finally to assign to non-audited taxpayers the same\(^{10}\) tax-evasion of the audited ones. As the assessment does not necessarily leads to ascertain the presence of amounts of tax evasion, not to all the taxpayers matched is assigned a level of tax evasion. Non audited individuals matched to audited for which there was no indication of evasion, are imputed of an amount of zero tax-evasion.

Data
The empirical analysis has been performed for the 2007 and 2008 fiscal years. In order to apply top-down and bottom-up approaches, different sources of data have been employed. First, top-down estimates requires national accounts data, provided by the National Statistics Institute (ISTAT), to estimate the potential tax base (BIT). Then, fiscal and administrative data on individual and companies’ income tax returns gathered from Tax Administration Data Base (‘Anagrafe Tributaria’) have been used to construct the BID for IRAP tax. Data on the cost of labor from ISTAT have also been used to decompose the IRAP base (declared and undeclared) into wages and gross profits.

As for the bottom-up approach, results from fiscal audits have been used. A special explanation of these data is needed due to the specificity of the activity of the Italian Revenue Agency. In fact, among all the possible fiscal controls, on-desk audits (namely, ‘accertamenti unificati’) represents surely the most complete form of control on a taxpayer’s activity. Due to an unavoidable delay in the availability of declared information from taxpayers, there is a lag between the fiscal (audited) year and the year in which the control is performed. Moreover, considering that it’s no more possible to control a fiscal year after five year, we consider that in order to have complete information on a specific audited tax year, data for five or six years of activity are needed, as shown in table 1. We use data on fiscal audits performed from 2009 to 2011 on 2007 and 2008 fiscal years. As our data are partial, results should be considered very preliminary.

Table 1: Audited fiscal year and year of control

<table>
<thead>
<tr>
<th>Year of control</th>
<th>(Audited) fiscal year</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>x</td>
</tr>
<tr>
<td>2007</td>
<td>x</td>
</tr>
<tr>
<td>2008</td>
<td>x</td>
</tr>
<tr>
<td>2009</td>
<td>x</td>
</tr>
<tr>
<td>2010</td>
<td>x</td>
</tr>
<tr>
<td>2011</td>
<td>x</td>
</tr>
<tr>
<td>2012</td>
<td>x</td>
</tr>
<tr>
<td>2013</td>
<td>x</td>
</tr>
</tbody>
</table>

Two databases of fiscal audits have been used to perform the bottom-up computations. To derive IRPEF and IRES base from gross profits, we use data on total audited taxpayers out of a population of more than 5 million firms, on average for 2007 and 2008.

In order to estimate individual income tax evasion (IRPEF) of self-employees and small size companies, we perform a propensity score matching analysis on a random stratified (by region and sector of economic activity) sample of 34,977 (17,626 audited and 17,351 not audited) firms selected from the population of taxpayers who completed income tax profile for 2007 and 2008 fiscal years.

\(^{10}\) After applying a non parametric procedure to control for outliers.
Main results
Due to the novelty of the approach and the availability of data, results of empirical analysis should be treated as very preliminary and should not be quoted.

Decomposition of IRAP BIND
Figure 3 shows the regional distribution of the 2007-2008 average total IRAP BIND on the potential tax base (BIT). Such ratio can be considered as a measure of the propensity in total base gap (Marigliani-Pisani, 2014). Results show that southern Italian regions have the most of propensity to evade tax base, with national average around 40%.

Figure 3 Ratio of Average (total) IRAP BIND on average BIT, 2007-2008.

As one of the aim of this paper, the disaggregation of the (total) gap in IRAP tax base has been possible using bottom-up techniques to estimates a measure of the intensity of evasion in tax base for IRES and IRPEF taxes, calculated as the ratio between the evaded tax base (BIND) on the declared one.

Using the intensity of BIND over BID derived from data on fiscal audits, it has been possible to divide the total IRAP BIND on gross profits into gap of IRES and IRPEF base, having subtracted the component of the cost of labor. Table 2 presents results on the relative weights of these two taxes on total undeclared gross profits. For both fiscal years analyzed, avoided tax base on declared base of medium-size and limited companies subjected to IRES (BIND IRES) has a relative higher weight on the total gross profits with respect to small firms and self-employees subjected to IRPEF (BIND IRPEF). It should be noted that IRES is due also by small and medium-sized firms with a simpler structure of business (namely, ‘società di persone’). These benefit from limited liability and thus are due to pay IRES but they can be much more similar in non-compliant behavior to self-employees than to big companies, considering also the features of the Italian economy. This can be true especially in regions where the number of small firms is relatively higher than big companies, as in the southern regions of Italy. Finally, the weight of tax evasion of IRES increases from 2007 to 2008, especially in regions where the number of bigger companies is relatively higher.
Table 2 Decomposition of Gross Profits

<table>
<thead>
<tr>
<th>Region</th>
<th>IRES 2007 (%)</th>
<th>IRES 2008 (%)</th>
<th>IRPEF 2007 (%)</th>
<th>IRPEF 2008 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIEMONTE</td>
<td>59</td>
<td>70</td>
<td>41</td>
<td>30</td>
</tr>
<tr>
<td>VALLE D’AOSTA</td>
<td>67</td>
<td>67</td>
<td>33</td>
<td>33</td>
</tr>
<tr>
<td>LOMBARDA</td>
<td>69</td>
<td>70</td>
<td>31</td>
<td>30</td>
</tr>
<tr>
<td>TRENTINO-ALTO ADIGE</td>
<td>53</td>
<td>55</td>
<td>47</td>
<td>45</td>
</tr>
<tr>
<td>VENETO</td>
<td>67</td>
<td>73</td>
<td>33</td>
<td>27</td>
</tr>
<tr>
<td>FRIULI-VENEZIA GIULIA</td>
<td>71</td>
<td>69</td>
<td>29</td>
<td>31</td>
</tr>
<tr>
<td>LIGURIA</td>
<td>67</td>
<td>71</td>
<td>33</td>
<td>29</td>
</tr>
<tr>
<td>EMILIA-ROMAGNA</td>
<td>69</td>
<td>69</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>TOSCANA</td>
<td>64</td>
<td>69</td>
<td>36</td>
<td>31</td>
</tr>
<tr>
<td>UMBRIA</td>
<td>65</td>
<td>64</td>
<td>35</td>
<td>36</td>
</tr>
<tr>
<td>MARCHE</td>
<td>63</td>
<td>65</td>
<td>37</td>
<td>35</td>
</tr>
<tr>
<td>LAZIO</td>
<td>69</td>
<td>73</td>
<td>31</td>
<td>27</td>
</tr>
<tr>
<td>ABRUZZO</td>
<td>67</td>
<td>73</td>
<td>33</td>
<td>27</td>
</tr>
<tr>
<td>MOLISE</td>
<td>68</td>
<td>76</td>
<td>32</td>
<td>24</td>
</tr>
<tr>
<td>CAMPANIA</td>
<td>62</td>
<td>65</td>
<td>38</td>
<td>35</td>
</tr>
<tr>
<td>PUGLIA</td>
<td>61</td>
<td>67</td>
<td>39</td>
<td>33</td>
</tr>
<tr>
<td>BASILICATA</td>
<td>61</td>
<td>63</td>
<td>39</td>
<td>37</td>
</tr>
<tr>
<td>CALABRIA</td>
<td>64</td>
<td>69</td>
<td>36</td>
<td>31</td>
</tr>
<tr>
<td>SICILIA</td>
<td>69</td>
<td>62</td>
<td>31</td>
<td>38</td>
</tr>
<tr>
<td>SARDEGNA</td>
<td>62</td>
<td>64</td>
<td>38</td>
<td>36</td>
</tr>
<tr>
<td>ITALY</td>
<td>64</td>
<td>67</td>
<td>36</td>
<td>33</td>
</tr>
</tbody>
</table>

Figure 4 shows the propensity in base gap for IRES and IRPEF as the ratio between undeclared base (BIND) over potential base (BIT).

Figure 4 Disaggregated Ratio of Average BIND on average BIT by macro regions, 2007-2008.

![Bar chart showing the percentage of base gap by macro regions.]

Given that the IRES has a flat statutory tax rate, by applying it the statutory tax rates to the IRES BIND, we obtain the IRES tax gap. In order to perform the same estimation on IRPEF it is necessary to further analyze the distribution of tax rate.
Estimation of Tax evasion of self-employees through propensity score analysis

Results of (1) are reported in Table 3 and are employed to calculate the propensity scores using nearest neighbor matching criterion (Shahidur et al. 2010). The scores for each individual, \( p_i \), are equal to:

\[
p_i = \Phi(x_i) = \text{Prob}(Y_i = 1 | x_i)
\]

Where \( \mathbf{X} \) is the vector of covariates used in the propensity model.

| Variables                      | Estimate | \( Pr(>|z|) \) |
|-------------------------------|----------|---------------|
| Intercept                     | -0.2545  | 0.0000        |
| Labor cost (amount of)        | -0.0001  | 0.0633        |
| Luxury cars (number of)       | 0.7229   | 0.0000        |
| Electricity consumptions (number of) | 0.1072    | 0.0000        |
| Net tax (amount of)           | -0.0016  | 0.0000        |
| Fac_Coherence (in the auditing scheme) | -0.2467    | 0.0000        |
| Fac_Not present (in the auditing scheme) | -0.0727    | 0.0004        |
| Fac_Congruence                | -0.2269  | 0.0000        |
| Employees (number of)         | 0.0613   | 0.0000        |
| Vat numbers (number of)       | 0.0825   | 0.0000        |
| Taxable income (amount of)    | 0.0007   | 0.0000        |
| Fac_Islands (macro region)    | -0.0523  | 0.0475        |
| Fac_North East (macro region) | -0.0213  | 0.3349        |
| Fac_North West (macro region) | 0.0118   | 0.5636        |
| Fac_South (macro region)      | -0.0251  | 0.2515        |
| Fac_Gender                    | 0.0930   | 0.0000        |
| Fac_Services (economic Sector) | 0.0300    | 0.2614        |
| Fac_Trade (economic Sector)   | -0.0009  | 0.9747        |
| Fac_Constructions (economic Sector) | -0.0741    | 0.0187        |
| Fac_Industry (economic Sector) | -0.0836    | 0.0222        |
| Activity's value (amount of)  | 0.0002   | 0.0009        |
| Vat credits (amount of)       | 0.0008   | 0.0000        |
| Vat debits (amount of)        | 0.0011   | 0.0000        |
| Passivity's value (amount of) | 0.0003   | 0.0000        |
| Added value (amount of)       | 0.0005   | 0.0204        |
| Taxable total base (amount of) | -0.0002   | 0.0528        |

Estimated coefficients are coherent with expected results. Those related to labor costs, net taxes and total taxable base reported in the individual income tax returns are significant and negatively linked with the probability to be audited, while amounts of assets, VAT credits, VAT debits, liabilities, value added and taxable income have all a positive impact on the probability to be audited. In addition, property of luxury cars, electricity consumptions and the count of VAT registration numbers have a positive impact too. A
dummy variable to indicate the gender of taxpayer has also been included and is significant and positive as to indicate a higher probability to be audited for male entrepreneurs with respect to female ones.

Industry and constructions are economic sectors in which taxpayers have lower probability to be audited than other sectors. Finally as for the accountability to a specific audit scheme (‘studi di settore’) results appear to indicate that the consistency and congruence criteria reduce the probability of being audited. Even the non-adherence to the scheme (that in Italy is voluntary and not mandatory) as respect to the incongruity allows taxpayers to reduce the probability to be audited.

In order to evaluate the model’s goodness of fit, the following statistical indices have been used: percentage of correct forecast audited (PCF):

\[ \text{Prob}(p_i > 0.5| R_i=1 \cap p_i \leq 0.5| R_i=0) \]

recall:

\[ \text{Prob}(p_i > 0.5| R_i=1) \]

and the log-likelihood of the fitted model (llh). Results are reported in Table 4.

<table>
<thead>
<tr>
<th>PCF</th>
<th>Recall</th>
<th>llh</th>
</tr>
</thead>
<tbody>
<tr>
<td>94.91%</td>
<td>77.64%</td>
<td>-23,130.51</td>
</tr>
</tbody>
</table>

Finally, under the assumption of conditional independence and a sizable overlap in propensity scores between participants and matched nonparticipants, the average treatment effect (ATE) in terms of declared gross income has been calculated. It is the mean difference in outcomes over the common support, weighting the comparison units by the propensity score distribution of participants registered in the sample and it is equal to 3139 euro with a standard error equal to 1089 (significant at 5% level). It indicates that on average, audited taxpayers declare a gross income of 3139 euro lower than non-audited ones.

Imputed values of self-employees tax evasion obtained using PSM analysis allow us to decompose the total gross profits tax evasion into three components related to different taxpayers: self-employees (SE) and small individual companies (SC) (due to pay IRPEF) and limited companies (LTD) (due to pay IRES). In particular, tax evasion of small individual companies is obtained deducting the aggregated post-stratified sample values of self-employees from the total bottom up estimation of IRPEF tax evasion (as shown in figures 1 and 2).

National undeclared total value added (BIND IRAP) and regional gross profits tax-evasion distributions are reported in Figure 5 and Table 5.

Figure 5 clearly shows the effects of tax evasion on primary distribution of income. In fact, considering that on average for 2007-2009, the weight of remuneration of labor over declared base is around 57% as in Braiotta et al (2013), we conclude that evasion operates a redistribution of benefits in favor of profits and entrepreneurs and against labor and employees. This results should also be considered in light of the fact that the choice of working in the black labor market is not always autonomous but induced by labor market features and the environment in which employees operate.

Studi di settore are based on a particular interaction between the tax agency and taxpayers, where the agency unveils only part of the information used to develop its audit rule. For details on the effects of this scheme on compliance, see Santoro and Fiorio (2011).
Figure 5 Italy’s Undeclared total Value Added (BIND IRAP) components for 2007.

Table 5 Regional gross profits tax-evasion components

<table>
<thead>
<tr>
<th>Region</th>
<th>IRES Limited Companies (LTD)</th>
<th>IRPEF Small Companies (SC)</th>
<th>Self-Employees (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIEMONTE</td>
<td>38.84%</td>
<td>28.43%</td>
<td>32.73%</td>
</tr>
<tr>
<td>VALLE D’AOSTA</td>
<td>37.01%</td>
<td>26.25%</td>
<td>36.74%</td>
</tr>
<tr>
<td>LOMBARDIA</td>
<td>54.05%</td>
<td>19.06%</td>
<td>26.89%</td>
</tr>
<tr>
<td>TRENTINO ALTO ADIGE</td>
<td>41.17%</td>
<td>29.56%</td>
<td>29.27%</td>
</tr>
<tr>
<td>VENETO</td>
<td>45.61%</td>
<td>26.40%</td>
<td>27.99%</td>
</tr>
<tr>
<td>FRIULI VENEZIA GIULIA</td>
<td>46.05%</td>
<td>28.76%</td>
<td>25.19%</td>
</tr>
<tr>
<td>LIGURIA</td>
<td>41.90%</td>
<td>27.44%</td>
<td>30.65%</td>
</tr>
<tr>
<td>EMILIA ROMAGNA</td>
<td>47.96%</td>
<td>25.58%</td>
<td>26.46%</td>
</tr>
<tr>
<td>TOSCANA</td>
<td>38.43%</td>
<td>27.37%</td>
<td>34.21%</td>
</tr>
<tr>
<td>MARCHE</td>
<td>39.19%</td>
<td>27.36%</td>
<td>33.45%</td>
</tr>
<tr>
<td>UMBRIA</td>
<td>39.86%</td>
<td>26.76%</td>
<td>33.37%</td>
</tr>
<tr>
<td>LAZIO</td>
<td>55.98%</td>
<td>18.31%</td>
<td>25.71%</td>
</tr>
<tr>
<td>ABRUZZO</td>
<td>37.17%</td>
<td>26.19%</td>
<td>36.64%</td>
</tr>
<tr>
<td>MOLISE</td>
<td>32.24%</td>
<td>33.28%</td>
<td>34.48%</td>
</tr>
<tr>
<td>CAMPANIA</td>
<td>40.90%</td>
<td>22.74%</td>
<td>36.36%</td>
</tr>
<tr>
<td>PUGLIA</td>
<td>32.74%</td>
<td>29.74%</td>
<td>37.51%</td>
</tr>
<tr>
<td>BASILICATA</td>
<td>31.91%</td>
<td>35.53%</td>
<td>32.56%</td>
</tr>
<tr>
<td>CALABRIA</td>
<td>30.57%</td>
<td>30.43%</td>
<td>39.00%</td>
</tr>
<tr>
<td>SICILIA</td>
<td>35.81%</td>
<td>20.51%</td>
<td>43.68%</td>
</tr>
<tr>
<td>SARDEGNA</td>
<td>32.88%</td>
<td>22.44%</td>
<td>44.68%</td>
</tr>
<tr>
<td>ITALY</td>
<td><strong>45.01%</strong></td>
<td><strong>23.85%</strong></td>
<td><strong>31.14%</strong></td>
</tr>
<tr>
<td>NORTH WEST</td>
<td>49.24%</td>
<td>22.07%</td>
<td>28.68%</td>
</tr>
<tr>
<td>NORTH EAST</td>
<td>46.24%</td>
<td>26.53%</td>
<td>27.22%</td>
</tr>
<tr>
<td>CENTER</td>
<td>47.87%</td>
<td>22.55%</td>
<td>29.58%</td>
</tr>
<tr>
<td>SOUTH</td>
<td>36.00%</td>
<td>25.01%</td>
<td>38.99%</td>
</tr>
</tbody>
</table>
Some extensions

As stated in the introduction, the influence of tax rates on compliance behavior is a central question for policy makers in evaluating the effects of tax reforms on public revenues. Starting from the pioneering model of Allingham and Sandmo (1972), many authors have developed studies regarding the relationship between tax rates and evasion with ambiguous results. On one hand, some find a positive direct causality between tax rates and tax evasion estimating also the elasticity of this relationship (Clotfelter, 1983); on the other, some other studies contradict the first ones arguing that tax rates have a negative effect on evasion (Feinstein 1991). The measurement of the effects of tax rate on tax evasion are outside the purpose of this paper. Nevertheless, results from PSM analysis on the tax return’s microdata of self-employees allow us to exert some empirical considerations with regard to this issue. In Figure 6 the percentage of estimated per-capita tax evasion on national average estimated tax evasion (TEpc) per actual tax rates (after deductions) is reported. It shows how TEpc generally increases with actual tax rates confirming the positive relationship between tax rates and evasion.

Figure 6 Actual tax rates and Incidence of tax Evasion

Figure 7 shows the relationship between the TEpc and legal tax rates and it appears to confirm the increasing link between tax evasion and tax rates, even with different patterns and elasticities depending on the level of tax rate. Figure 8 plots the actual, the average (considering tax evasion) and the marginal tax rates (percentage taken from next amount of taxable income once imputed amount of evasion) for each tax bracket. The majority of evasion is concentrated among individuals that have a declared level of income till the fourth income bracket. For high level of income, there is a small convenience to evade as the difference between average (including evasion) and actual tax rate is low.

12 We should remind that the distribution of taxpayers among different tax rates is based on the level of declared income after deductions. In order to better exploit the relationship between evasion, we should support the analysis contained in figures 6 and 7 with the that of gross declared income (before deductions).
In terms of inequality analysis, results of PSM analysis allow us also to investigate the impact of tax evasion on the secondary (horizontal) income distribution. In order to reach this goal, we compare the Lorenz curves of the distribution of declared income $y$ (with evasion) and of after tax-audited income, $z = y + t^*$ (without evasion), where $t^* \geq 0$ is the amount of marginal taxable income imputed after tax-audit. Figures 9 shows that the existence of tax evasion generates a distortion in equality terms, since the distribution $z$ dominates in the Lorenz sense the distribution $y$\textsuperscript{13}.

\textsuperscript{13} Similar results are obtained by Serén and Panadés (2005) using Spanish Personal Income Tax data.
Corresponding values for the Gini index (reported in Table 6) confirm the existence of a negative bias effect, in terms of inequality, of the tax-evasion on income distribution. Despite the strong inequality of both distributions, \( z \) is more equally distributed than \( y \) since the Gini index is clearly inferior.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Gini Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>( z )</td>
<td>0.602</td>
</tr>
<tr>
<td>( y )</td>
<td>0.634</td>
</tr>
</tbody>
</table>

**Table 6** Gini indexes of income distributions

Figure 9 Lorenz Curves: Declared income distribution with evasion and without evasion

Concluding remarks

The aim of the paper is to integrate top-down and bottom-up approaches to derive more disaggregated figures of total tax gap into single taxes and for different taxpayers. This decomposition helps deriving conclusion on the distribution of primary income between labor cost and gross profits, before and after considering tax evasion and on secondary distribution of income when we further divide personal income tax for different types of taxpayers.

The principal conclusions of our analysis can be summarize as follows:

- Tax evasion operates a redistribution in favor of profits, as the weight of total undeclared profits is much higher than that of labor;
- The incidence of gap in IRES tax base is higher than IRPEF tax base indicating that the juridical organization of firms and their relative size leads to greater opportunity to avoid or limit tax burden;
Comparing our results on the ratio of BIND over BID with previous similar studies, as Braiotta et al. (2013), it is confirmed that tax avoidance shifts benefits from the component of remuneration of labor to that of profits;

Between 2007 and 2008, the propensity to gap base for IRPEF and IRES increases with different regional patterns: while for IRES to the contribution to the increase is positive for all the macroregions, north western regions experiences a decrease in the incidence of IRPEF evasion;

Gross income tax evasion components are homogeneously distributed between Italian regions

Moreover, even though this work is only a preliminary step for further and more detailed micro-analyses (topics for future research), some interesting issues have raised regarding the effect of tax rates on compliance and the effect of tax evasion on the income distribution.

An excessive tax burden may be counterproductive in terms of compliance because it seems that evasion increases with the tax burden and tax rates;

Results confirm the existence of a negative bias effect of tax evasion on income distribution in a progressive tax system since self-employees’ income without evasion is more equally distributed than including evasion.
References


Giovannini, E. (2010), "Indagine conoscitiva su taluni fenomeni distorsivi del mercato del lavoro (lavoro nero, caporalato e sfruttamento della manodopera straniera)”. Audizione alla Camera dei Deputati, XI Commissione permanente "Lavoro pubblico e privato”.


