UNIVERSITY OF Southampton
School of Social Sciences

School of Social Sciences
Economics Division
University of Southampton
Southampton SO17 1BJ, UK

Discussion Papers in
Economics and Econometrics

Too Low to be True: The Use of Minimum Thresholds to Fight Tax Evasion

Mirco Tonin

No. 1018

This paper is available on our website
http://www.southampton.ac.uk/socsci/economics/research/papers

ISSN 0966-4246

Too low to be true: the use of minimum thresholds to fight tax evasion

Mirco Tonin

Economics Division,
School of Social Sciences,
University of Southampton,
Southampton SO17 1BJ,
United Kingdom
m.tonin@soton.ac.uk

September 28, 2010

1 Introduction

The enforcement of compliance with tax regulation is a complex task.[1] This is particularly the case when the administrative capacity of the tax authority is low, as is often the case in developing and transition countries (Bahl and Bird, 2008). This paper draws on some international experiences in fighting tax evasion to identify tools that can be used to reduce underreporting by employed labor, small and medium enterprises, self-employed, and professionals.

I will analyze in some details two policies: the Italian "Business Sector Analysis", which targets small and medium enterprises, self-employed, and professionals and the Bulgarian "Minimum Social Insurance Thresholds", which is aimed at employees. In the Italian system, taxpayers reporting revenues below a certain threshold are subject to more intense scrutiny by the tax authority, while in the Bulgarian case, social security contributions paid by firms and employees should be above a certain threshold. In both cases, the thresholds depend on quantities that are correlated with income and are more easily observed. As these thresholds represent "indirect means to ascertain tax liability, which differ from the usual rules based on the taxpayer's accounts" (Thuronyi, 1996), they belong to the category of presumptive taxation methods, with the Italian system being rebuttable and the Bulgarian one irrebutable.

In this paper, I first formally model the impact of minimum thresholds by explicitly taking into account the low administrative capacity characterizing many developing and transition countries. In particular, I depart from the assumption of perfect detection in case of an audit that is standard in the literature. This assumption is problematic for developed countries (see the studies by Feinstein, 1991, and Erard, 1997, for the US case) and is clearly untenable for developing and transition ones. The model shows that the introduction of a threshold creates a spike and a "missing

middle" in the distribution of declared incomes and highlights under which conditions introducing a threshold is likely to increase net revenues for the tax authority.

The use of thresholds is of particular interest for two set of reasons. From a normative perspective, minimum thresholds have been found to describe the optimal auditing strategy in several settings. Reingaum and Wilde (1985) compare a "standard random audit policy" to "an 'audit cutoff' policy, in which an agent triggers an audit if reported income is 'too low'" and find the latter to weakly dominate the former. Similar results are also found in Sanchez and Sobel (1993) and Macho-Stadler and Perez-Castrillo (1997). A recent contribution is Bigio and Zilberman (2010) who study the optimal monitoring of self-employed entrepreneurs, allowing for the tax authority to condition the monitoring strategy on labor input. They find that it is optimal for the tax authority to calculate inputted income as a function of labor and then audit entrepreneurs who report below inputted income. A second reason why thresholds are of interest is that, from an administrative point of view, minimum thresholds are, at least potentially, easy to understand for the taxpayer and to implement for the tax authority. As Bird (2004) notices "[t]he best tax policy in the world is worth little if it cannot be implemented effectively" and this constraint is particularly binding in the case of developing and transition countries, so that "an essential precondition for the reform of tax administration is to simplify the tax system in order to ensure that it can be applied effectively in the generally low-compliance contexts". Thus, the potential simplicity of minimum thresholds makes them relevant from a policy perspective and the description of the Italian and Bulgarian systems is intended to provide some guidance for policymakers concerning their implementation.

The rest of the paper is organized as follows. In the next section, I will introduce a simple model that allows to study the impact of minimum thresholds in an environ-

ment characterized by low administrative capacity. Then, in sections 3 and 4, I will analyze the Italian "Business Sector Analysis" and the Bulgarian "Minimum Social Insurance Thresholds", describing how they are elaborated and applied and appraising their applicability as an instrument to fight undeclared work and tax evasion. Finally, the last section of the paper concludes and provides policy recommendations.

2 A model with imperfect detection

I first consider an environment without a threshold and then investigate the impact of introducing one. The analysis is conducted for one specific audit class, defined according to some criteria like sector or occupation, with many different audit classes potentially present in the economy as a whole (Scotchmer, 1987). The threshold I consider is a rebuttable one, close to the Italian case, where taxpayers declaring below it are subject to higher scrutiny. I analyze the irrebutable case, close to the Bulgarian case, where taxpayers are not allowed to declare less than a minimum threshold in Tonin (2007). The results are similar.

Consider a taxpayer, either an individual or a firm, with exogenous income y, who faces a tax rate $t \in (0,1)$. His declaration to the tax authority is denoted by $x \in [0,y]$. The tax authority may conduct an audit to find out whether he complies with fiscal regulation. I assume, for the moment being, that there is one exogenously given probability of an audit, $\gamma \in [0,1]$. A fine proportional to the amount evaded is imposed in case tax evasion is detected. However, the fact that an audit is performed does not imply that the authority discovers with certainty the true tax liability. It may instead find evidence to impute an income $\hat{y} \in [0,y]$, where y is the true income.

I assume that $\hat{y}$ is distributed over the support $[0,y]$ according to pdf $h(\cdot)$ and cdf $H(\cdot)$, so that $H(0)=0$ and $H(y)=1$, and $H(\cdot)$ does not depend on x.[2] To simplify

the discussion, I assume that $h(\cdot) > 0$ within the support, so that $H(\cdot)$ is invertible within $[0, y]$.

Given a declaration of x and collected evidence of a true tax liability of $\hat{y}$, the tax authority imposes, in case $\hat{y} > x$, the payment of $\theta t\,(\hat{y} - x)$, consisting of taxes plus an additional fine proportional to the assessed tax evasion, thus $\theta > 1$. This specification of the fine, proportional to the amount of tax evaded, follows Yitzhaki (1974). In case $\hat{y} \leq x$, the tax authority cannot prove any tax evasion, so no fine is imposed.[3] Given a true income y and a reported one x, the expected fine in case of auditing, f, is

$$f = t\theta \int_{x}^{y} (\hat{y} - x) h(\hat{y}) d\hat{y}. \tag{1}$$

The taxpayer is risk-neutral and maximizes expected net income, given by

$$I = y - \gamma f - tx, \tag{2}$$

where tx are payments due to voluntary compliance and γf payments due to enforcement. Therefore, the optimal declaration is given by

$$x^* \quad s.t. \quad \max_{x \in [0,y]} y - \gamma f - tx. \tag{3}$$

After substituting 1 into 3, the first-order condition is

$$H(x^*) = 1 - \frac{1}{\gamma\theta} \Longleftrightarrow x^* = H^{-1}\left(1 - \frac{1}{\gamma\theta}\right). \tag{4}$$

The second-order condition, $-t\gamma\theta h(x) < 0$, is always satisfied. The boundary condition $x \leq y$ is always satisfied. The condition $x \geq 0$ implies that full evasion will take place, i.e. $x = 0$, when enforcement is very weak, i.e. $\gamma\theta \leq 1$. To simplify

the notation, the two enforcement parameters are summarized by $\alpha \equiv 1/(\gamma\theta)$. To summarize, the solution to the reporting problem is given by

$$x^* = \begin{cases} H^{-1}(1-\alpha) & \text{if} \quad \alpha < 1 \\ 0 & \text{if} \quad \alpha \geq 1 \end{cases}. \tag{5}$$

As $\partial\alpha/\partial\gamma < 0$ and $\partial\alpha/\partial\theta < 0$, in an interior solution the fraction of income that is evaded decreases as enforcement improves. The equilibrium fine, f^*, is given by substituting 5 into 1. Expected net income in equilibrium, I^*, is then given by substituting x^* and f^* into 2.

To obtain a closed form solution, from now on I will assume $h(\cdot)$ to be uniform in the support $[0, y]$, i.e. $\hat{y} \sim U_{[0,y]}$. The expression for the expected fine becomes

$$\gamma f = t(y-x)^2/(2\alpha y). \tag{6}$$

Thus, the cost of evasion is quadratic in the amount of evasion, $y - x$. The optimal reporting behavior given by 5 becomes

$$x^* = \begin{cases} (1-\alpha)\, y & \text{if} \quad \alpha < 1 \\ 0 & \text{if} \quad \alpha \geq 1 \end{cases}. \tag{7}$$

Thus, the model implies that a fraction of the true tax liability that depends on the enforcement parameters is revealed to the fiscal authorities. Substituting 7 into 6, the expected fine is given in equilibrium by

$$\gamma f^* = \begin{cases} yt\alpha/2 & \text{if} \quad \alpha < 1 \\ yt/(2\alpha) & \text{if} \quad \alpha \geq 1 \end{cases} \tag{8}$$

and, substituting 7 and 8 into 2, I get the equilibrium expected net income

$$I^* = \begin{cases} y(1-t) + \alpha y t/2 & \text{if} \quad \alpha < 1 \\ y\left[1 - t/\left(2\alpha\right)\right] & \text{if} \quad \alpha \geq 1 \end{cases}. \tag{9}$$

Given the detection technology, the expected fraction of unreported tax liability, $y - x^*$, that is discovered in case of auditing is

$$\frac{1}{(y-x^*)} \int_x^y (\hat{y} - x^*) h(\hat{y}) d\hat{y} = \alpha/2, \tag{10}$$

i.e. a fraction corresponding to half the ratio of evaded income over true product. Thus, it is relatively easy to get away with tax evasion: for example, in an economy where 30% of the income is concealed, only 15% of evasion is, on average, detected in case of auditing. Consistently with the low administrative capacity that characterizes many developing and transition countries, the assumption is thus that an audit is quite ineffective.

2.1 The effect of a threshold

I now consider the impact of increasing the probability of an audit in case of a declaration below a certain threshold, $\bar{x}$, so that the probability of an audit is now given by

$$\begin{cases} \gamma_H & \text{if} \quad x < \bar{x} \\ \gamma_L & \text{if} \quad x \geq \bar{x} \end{cases} \quad \text{with } \gamma_H > \gamma_L. \tag{11}$$

I can then define $\alpha_H \equiv 1/\left(\gamma_H \theta\right)$ and $\alpha_L \equiv 1/\left(\gamma_L \theta\right)$, with $\alpha_L > \alpha_H$. To simplify the discussion, I assume that $\alpha_L < 1$. Therefore, even with a low audit probability, deterrence is strong enough not to have full evasion. Notice that for any given y

and x, expected income as given by 2, is greater in an environment with a low audit probability than in an environment with a high audit probability. Following 7, I define $x_H = (1-\alpha_H)\, y$ and $x_L = (1-\alpha_L)\, y$, with $x_H > x_L$ and, following 9, $I_H^* = y(1-t) + \alpha_H yt/2$ and $I_L^* = y(1-t) + \alpha_L yt/2$.

I can now characterize the optimal behavior for a taxpayer in an environment with a minimum threshold. For taxpayers with high income, i.e. with $y > \bar{x}/(1-\alpha_L)$, x_L is above the threshold and, therefore, the low audit probability does indeed apply if x_L is declared. Then, the optimal declaration is given by $x^* = x_L$. For taxpayers with a lower income x_L is instead unattainable, as it is below the threshold and therefore the high audit probability applies in case x_L is declared. Then, there are two possible courses of action: one is to declare exactly the threshold and face a low audit probability; the other is to face a high audit probability and declare the best possible amount in such an environment, x_H. I can exclude all other possibilities as declaring at the threshold dominates any higher declaration for these taxpayers and declaring x_H dominates any other declaration for which a high audit probability applies. The level of gross income at which the taxpayer is indifferent between declaring $\bar{x}$ and x_H is given by equating I_H to the level of net income when the declaration is $\bar{x}$, calculated by replacing $\bar{x}$ in 1 and 2. The taxpayer is indifferent when gross income is $\bar{y} = \bar{x}/\left(1-\alpha_L + \sqrt[2]{\alpha_L(\alpha_L-\alpha_H)}\right)$. Thus, for taxpayers with intermediate income, i.e. $\bar{x}/(1-\alpha_L) \geq y \geq \bar{x}/\left(1-\alpha_L + \sqrt[2]{\alpha_L(\alpha_L-\alpha_H)}\right)$, it is optimal to declare exactly at the threshold, i.e. $x^* = \bar{x}$. For taxpayers with even lower income, $y < \bar{x}/\left(1-\alpha_L + \sqrt[2]{\alpha_L(\alpha_L-\alpha_H)}\right)$, it is instead optimal to declare below the threshold, with $x^* = x_H$.

Now, I investigate the implications of the optimal individual behavior outlined above for the distribution of declared income. Suppose that taxable income, y, is distributed in the population according to pdf $g(y)$ on the support $[y_{\min}, y_{\max}]$, with

$y_{\min} > 0$. Then, the distribution of declared income x in an environment without a threshold, g_x, is given by

$$g_x(x) = \begin{cases} g\left(\frac{x}{1-\alpha}\right) & \text{if} \quad y_{\min}(1-\alpha) < x < y_{\max}(1-\alpha) \\ 0 & \text{otherwise} \end{cases}. \tag{12}$$

as taxpayers with income y declare $(1-\alpha)\,y$ instead. The distribution of declared income when there is a threshold, $g_{\bar{x}}$, is

$$g_{\bar{x}}(x) = \begin{cases} g\left(\frac{x}{1-\alpha_H}\right) & \text{if } y_{\min}(1-\alpha_H) < x < \bar{y}(1-\alpha_H) \\ \int_{\bar{y}}^{\bar{x}/(1-\alpha_L)} g(y)\,dy & \text{if } x = \bar{x} \\ g\left(\frac{x}{1-\alpha_L}\right) & \text{if} \quad \bar{x} < x < y_{\max}(1-\alpha_L) \\ 0 & \text{otherwise} \end{cases}. \tag{13}$$

Notice that $(1-\alpha_H)\,\bar{y} < \bar{x}$, therefore there is no taxpayer declaring an income in the interval $[(1-\alpha_H)\,\bar{y}, \bar{x})$. The distribution of declared income when there is a threshold that triggers a higher audit probability is then characterized by a spike at the threshold, with taxpayers bunching there to avoid being subject to the higher audit probability, and by a "missing middle", with nobody declaring an income just below the threshold.

I can now study the budgetary implications of introducing a minimum threshold. Without a threshold, payments from each taxpayer, P, are given by the sum of voluntary compliance, $(1-\alpha)\,ty$, and enforcement, given by 8, $\alpha ty/2$, thus

$$P = (1-\alpha/2)\,ty, \tag{14}$$

and total fiscal revenues, R, are given by

$$R = \int_{y_{\min}}^{y_{\max}} (1 - \alpha/2)\, tyg\,(y)\, dy. \tag{15}$$

If the cost per audit is c, then the total cost for the tax authority, C, after normalizing the size of the population to 1, is $c\gamma$.

Suppose now that the tax authority introduces a threshold. Payments for taxpayers at the upper part of the income distribution, P_U, and at the bottom of the income distribution, P_B, are given by 14, with α replaced by α_L and α_H respectively. Payments by taxpayers with intermediate income, the ones declaring exactly the threshold, $P_{\bar{x}}$, are given by the sum of voluntary compliance, $t\bar{x}$, and enforcement, obtained by replacing $\bar{x}$ into 6 and taking into account that the low audit probability applies, $t\,(y - \bar{x})^2 / (2\alpha_L y)$. Total revenues are then given by

$$R_{\bar{x}} = \int_{y_{\min}}^{\bar{y}} P_B g\,(y)\, dy + \int_{\bar{y}}^{\bar{x}/(1-\alpha_L)} P_{\bar{x}} g\,(y)\, dy. + \int_{\bar{x}/(1-\alpha_L)}^{y_{\max}} P_U g\,(y)\, dy, \tag{16}$$

while costs are given by

$$C_{\bar{x}} = c\gamma_H \int_{y_{\min}}^{\bar{y}} g\,(y)\, dy + c\gamma_L \int_{\bar{y}}^{y_{\max}} g\,(y)\, dy. \tag{17}$$

To simplify the rest of the discussion, assume that $\gamma_L = \gamma$, so that the introduction of a threshold is actually an increase of the audit probability for taxpayers declaring below it, while the audit probability remains unchanged for those above the threshold. Then, the change in revenues due to the introduction of a threshold is given by

$$\Delta R = \int_{y_{\min}}^{\bar{y}} \frac{\alpha - \alpha_H}{2} tyg\,(y)\, dy + \int_{\bar{y}}^{\bar{x}/(1-\alpha)} \frac{[\bar{x} - (1-\alpha)\, y]^2}{2\alpha y} tg\,(y)\, dy, \tag{18}$$

while the change in costs is given by

$$\Delta C = c\left(\gamma_H - \gamma\right) \int_{y_{\min}}^{\bar{y}} g\left(y\right) dy. \tag{19}$$

The increase in revenues due to the introduction of a threshold derives from an increase in the declaration for those actually facing a higher probability of an audit, the first term in 18, and an increase in the declaration for those declaring at the threshold to avoid the higher probability of an audit, the second term in 18. While the former category of taxpayers entails a higher cost for the tax authority as audits are more frequent, the latter does not, as the probability of an audit remains unchanged. So, the introduction of a threshold is likely to increase net revenues for the tax authority, $R - C$, if it induces a large number of taxpayers to increase their compliance by declaring at the threshold. Whether this is the case depends on the position of the threshold within the gross income distribution and on the audit probabilities, in particular on the difference in audit probabilities when above and below the threshold, as $\bar{y}$ depends on $(\alpha_L - \alpha_H)$. On the other side, the increase in compliance due to taxpayers facing a higher audit probability is unlikely to increase net revenues, as this taxpayers are characterized by low gross income and therefore the increase in revenues is not going to be substantial, while the increase in the number of audits is costly.[4] The optimal choice of threshold and audit probabilities is outside the scope of this paper. However, the discussion above highlights some of the factors that need to be taken into account. In the rest of the paper, I will provide two examples of how thresholds are actually implemented.

3 The Italian Business Sector Analysis

Business Sector Analysis[5,6] (BSA thereafter, "studi di settore" in Italian) is an instrument used by the Italian tax authority to estimate revenues and compensations for small and medium enterprises, self-employed, and professionals. It has been first applied in 1998 and revised several times in subsequent years. BSA represents a hybrid between an auditing selection mechanism and a form of presumptive taxation (Santoro, 2006). BSA is an indirect mean to ascertain tax liability, mostly based on evidence that is deemed more reliable than the taxpayer's accounts. However, taxes are paid on reported profits, while BSA only estimates revenues. Moreover, there is no obligation for the taxpayer's declared revenues to match the estimate. A failure to do so only entails an increased probability of an audit and the taxpayer has the right to come forward with proper evidence to substantiate lower revenues.

In what follows, I will look at how a BSA is elaborated and applied. A critical appraisal of BSA as an instrument to fight undeclared work and tax evasion concludes this section.

3.1 The elaboration of Business Sector Analysis

The number of BSA has increased from 45 in 1998 to more than 200 in recent years. There is, for instance, a BSA applying to "Retail sale of flowers, plants and seeds via permanent or mobile stalls" and a BSA on "Tour guide activities" or "Dentistry". The elaboration of a BSA for a particular sector is based on a procedure consisting of several steps. A regular updating is necessary to ensure that a BSA captures the economic structure characterizing a sector. For this reason a "revision" of a BSA has to be produced at least every 4 years.

3.1.1 Data collection

The first step is to collect data through a survey of the population of interest. The characteristics of the survey are defined through consultations with the relevant trade association. Answering the survey is compulsory, but failure to do so is not sanctioned. The questionnaire developed in 2000 for the elaboration of the aforementioned BSA on retail sale of flowers, for instance, was sent to almost 27,000 businesses, but only 19.3% was sent back, of which almost 60% could not be used because of incomplete data or other issues, leaving only 2,168 questionnaires available for the analysis. The information collected regarded:

1. personnel (13 questions, e.g. number of part time employees)
2. characteristics of the business premises (25 questions, e.g. daily opening time, parking area reserved for customers, area used as storage)
3. means of transportation (5 questions, e.g. number of trucks between 3.5 and 12 tones)
4. products and customer base (24 questions, e.g. % of revenues realized through the sale of dried flowers or through the rental of ornamental plants)
5. supply channels (3 questions, e.g. % of purchases directly from producers)
6. miscellaneous data (14 questions, e.g. home delivery services, membership in delivery networks)

Beside these "structural" data, a set of accounting data was also collected (25 questions), mainly regarding inventory and costs.

3.1.2 Identification of sectorial clusters

The data collected are then used to divide the sector into homogenous groups of taxpayers, in the meaning of groups of firms, self-employed or professionals charac-

terized by the same organizational structure, operating in the same market segment, serving the same type of customers, using a similar business model.

From the dataset collected through the sectorial survey, the most significant variables are selected through Principal Component Analysis (PCA), a statistical technique used to reduce the dimensionality of a dataset by minimizing the loss of information. As the aim is to distinguish taxpayers according to their structural characteristics, the accounting data collected in the survey are not used at this stage.

Using these selected variables, the groups are then identified through Cluster Analysis, a statistical technique used to divide a dataset into subsets (clusters), so that the elements of each subset are similar to each other and dissimilar to elements of other clusters. For instance, in the aforementioned BSA regarding retail sale of flowers, 8 different clusters have been identified depending on location (operating within marketplaces, nearby cemeteries), sale point type (mobile stall, permanent stall, kiosk), and size.

3.1.3 Identification of geographical clusters

Given the importance of location for businesses, an attempt is made to classify the administrative units (municipalities, provinces and regions) into which the national territory is divided into homogenous groups according to their socioeconomic development. The aim is to capture both the differences in terms of final demand and the availability or not of business services, like logistic and credit. The variables used refer to:

1. educational achievement (3 variables, e.g. % of population with at least secondary degree)

2. wealth indicators (5 variables, e.g. per capita disposable income, bank deposits, cars above 2000cc)

3. business development indicators (6 variables, e.g. number of firms active in the transport sector every 100 inhabitants)

The cluster analysis applied to these variables has determined 5 geographical clusters. For instance, one group comprising mostly small municipalities in northern Italy is defined as "areas with high level of development, high educational achievement, and developed production structure".

For some sectors a separate clustering of the national territory is performed, accounting for their specific characteristics. For the trade sector, for instance, some additional variables concerning the structure of distribution channels are taken into account (8 variables, e.g. number and size of supermarkets per 1000 inhabitants), producing a classification of the national territory into 7 clusters (e.g. "areas with high level of development, developed industrial structure, and traditional distribution channels"). Other sectors with a distinctive geographical clustering are tourism and several manufacturing activities organized within so-called "industrial districts" (e.g. ceramic, gold, furniture), for which the concentration of similar activities within a small area is an important competitive factor.

Each municipality, province, and region is assigned to one cluster belonging to the general or to a sector-specific classification.

3.1.4 Estimation of the revenue function

To estimate the revenue function for each sectorial cluster, a group of taxpayers is selected among whose answering the survey. The aim is to exclude taxpayers with "anomalous" declarations, so that the estimate can better reflect the situation of a firm, self-employed or professional operating under "normal economic circumstances". Selection is based on "consistency indicators" (e.g. value added per employee, mark-up) and on the distribution within clusters of these indicators. So, for

instance, for the cluster "flower retailers with kiosk operating nearby cemeteries", firms with value added per employee in the lower 20th percentile or with a mark-up outside the interval determined by the 20th and 95th percentiles are excluded. Moreover, firms declaring total revenues lower than costs are excluded outright. These criteria are determined through "expert judgement".

Regression analysis is then applied to the selected sample of taxpayers considered "normal" to determine the relationship between revenues and the independent variables (structural and accounting). The independent variables to be included are determined through a stepwise selection procedure. Moreover, the effect of location may be accounted for by adding dummy variables for the geographical clusters and by interacting them with the most important structural or accounting variables.

The final result of the econometric analysis is a set of coefficients relating revenues to some independent variables. For the above mentioned cluster, for instance, the coefficient associated with the variable "number of family members working in the firm" is 5,192, meaning that estimated revenues will increase by that amount (in EUR) for any additional family member operating in the firm. The variable "cost of goods sold" has a coefficient 1.1509, "corrected" with an additional factor of 0.0689 if the firm is located in a municipality belonging to the trade-specific geographical cluster "areas with high level of development, developed industrial structure, and traditional distribution channels".

The tax authority can declare some sectors to be in a cyclical downturn and correct the estimates by a "cyclical correction factor". This can be automatic, applying directly to the function that estimates revenues, or non-automatic, in which case it can be applied by the local tax office to a taxpayer if properly motivated.

3.1.5 The application of Business Sector Analysis

Not all taxpayers are subject to a BSA. First of all, a BSA covering the sector in which a taxpayer operates must exist. Even if a BSA has been elaborated, taxpayers with revenues over 5,164,569 EUR are exempted. All in all, more than 4 million firms are subject to a BSA, representing in 2004 approximately 86% of all taxpayers. There are also other causes of exclusion, like starting or ceasing the activity in the fiscal year, having changed sector during the year or interrupted the activity due to renovation of business premises. The process of applying a BSA to a single taxpayer can be divided into several steps.

Determination of estimated revenues Taxpayers for which a BSA applies are required to fill a form for the communication of the required data to the tax authority within the same deadlines set forth for the income tax return. The failure to submit is penalized by a fine.

The first step to estimate revenues of a specific taxpayer is the determination of the cluster to which it belongs within the sector. This is done through discriminant analysis, a statistical technique used to classify objects into groups. A single firm may belong to more than one cluster, each with a given probability or weight. Then, the coefficients estimated for a given cluster are applied to the independent variables provided by the taxpayer, to determine both a point estimate and a confidence interval for estimated revenues. Only the lower threshold of this interval matters.

If a taxpayer belongs to more than one cluster, then estimated revenues are calculated as a weighted average of the estimates for each cluster.

Declaration Once the data in the BSA form are filled in, the point estimate and the lower threshold are calculated using computer software and are available to the

taxpayer before sending in the declaration form. If revenues arising from the taxpayer's accounts are at or above the point estimate, then the taxpayer is said to be "naturally consistent". If, however, accounting revenues are lower, the taxpayer can decide to increase the declared revenues to that level. In this case a taxpayer is said to be "consistent by adjustment". If the adjustment required to be consistent is above 10%, then a penalty rate of 3% applies. A taxpayer can also decide to increase the declared revenues to a level between the lower threshold and the point estimate, provided that proper motivation is given. Finally, a taxpayer can decide not to adjust and declare the (lower) accounting revenues. In this case a taxpayer is classified as "inconsistent".

Additional criteria to achieve consistency have been recently introduced, based on coherence of main economic indicators (e.g. value added per employee) with what is regarded as the "norm" in the sector. The aim is to discourage the manipulation of independent variables.

Auditing Taxpayers that are classified as inconsistent may be subject to an audit and invited to justify the reasons why their declared revenues fall below the estimated value. It is the taxpayer who has the burden of proof to establish the facts according to which the presumption derived from the BSA is not applicable to his or her specific circumstances. The tax administration provides several examples of acceptable reasons to justify the "inconsistency". One case is a serious illness or a maternity status that makes the normal business activity impossible. A firm documenting that it is working only with public institutions can also have its accounting revenues prevail over the estimated ones. Another example is the case of a shop experiencing a reduced business due to construction works nearby. More generally, a taxpayer can claim that the business activity is conducted in an "economically

marginal" or disadvantaged situation, such that the results of the BSA should not apply. All these circumstances have to be properly documented and in some cases should be certified by third parties, like tax assistance centres, tax practitioners and some trade associations. Also the reported "independent variables" may be subject to an audit.

3.2 An assessment of Business Sector Analysis

There hasn't been a systematic assessment of BSA. However, some data on their impact are available (see Convenevole et al., 2006) and are discussed below. The existing evidence suggests that BSA is effective if properly implemented. This section also includes some data on the elaboration costs and an evaluation of the pros and cons of BSA.

3.2.1 Impact

In 2004 almost 69% of taxpayers subject to a BSA resulted "naturally consistent", while the remaining 31% were "inconsistent". Of this, 47% decided to become "consistent by adjustment", while the remaining 53% did not adjust the declaration. In terms of the model, the former group represent the spike at the threshold. Among the firms who decided to adjust, the average adjustment per capita was of around 6,000 EUR, bringing a total increase in declared revenues of almost 3 billions EUR. This represents only 13% of the increase that would have been achieved if all firms decided to fully adjust. Consistently with the model, this suggests that firms decide to adjust if the difference between account revenues and estimated revenues is relatively small.

The 3 billions EUR figure may seem modest, but it represents only a part of the

total impact of BSA. The most likely impact of an established and effective BSA is indeed to change the firms' accounting behavior, by reducing underreporting during the year.

As mentioned above, there is a need to regularly update BSA. It is interesting to disaggregate the 2004 data depending on whether or not the relevant BSA had been recently updated. In the former case, 63% of the taxpayers resulted "naturally consistent", while in the latter case the figure was much higher at 74%. As already mentioned, this may be due to the fact that taxpayers change their accounting behavior during the year in consideration of the BSA, so that they do not need to perform an adjustment in the declaration. This "fine tuning" is easier for a BSA that is unchanged, but becomes more difficult when an updated version replaces the old one, as different independent variables and different coefficients may be used in the estimation. It may also be the case that an updated BSA better reflects the economic condition and structure of a sector, thus being able to "capture" more of the potential revenues. Among the "inconsistent" taxpayers, 51% decided to become "consistent by adjustment" when the BSA had been recently updated; otherwise the figure is a lower 41%. For the former category of taxpayers, the average adjustment per capita was of 6,143 EUR, representing a 17% of the requested adjustment, bringing a total increase in declared revenues of almost 1.9 billions EUR. In case of a BSA that had not been recently updated the corresponding figures were 6,667 EUR, corresponding to 10% of requested adjustment, bringing in 1.1 billions EUR additional revenues.

The importance of effective enforcement is underlined by looking at the data disaggregated along another dimension. In 2004, the rules regarding audit in case of inconsistency were tougher for firms using simplified bookkeeping than for firms using ordinary bookkeeping. In particular, being inconsistent in any given year was enough

to trigger an audit based on BSA for firms in the simplified regime, while firms in the ordinary one needed to be inconsistent for two years in any given three years period. This is no longer the case. However, when this rule was in place, the risk of an audit in case of inconsistency for a firm with ordinary accounting was much lower. This clearly had an impact on the amount of adjustment in case of inconsistency. In 2004 single entrepreneurs with simplified bookkeeping performed an adjustment equivalent to 33% of the required one on average, while the figure for single entrepreneurs with ordinary accounting is much lower at 12.8%. For partnerships the figures are 36% and 10.8% respectively. This suggests that BSA can be effective provided the "threat" of an audit in case of inconsistency is real.

3.2.2 Costs

BSA are elaborated and updated by SOSE or "Societa' per gli Studi di Settore S.p.A." (Corporation for Business Sector Analysis Inc.), a joint stock company participated by the Treasury and the Bank of Italy and active since 2002.

The costs associated with the elaboration and management of BSA amounted to around 7 millions EUR in 2004 and 2005, increasing to approximately 8.5 millions EUR in 2006, with labor costs accounting for around half of the total.

3.2.3 Pros and cons

The aim of BSA is to go beyond (possibly unreliable) accounting data and use easier to ascertain characteristics and their relationship with revenues to determine "expected revenues". For this to be feasible it is necessary that a set of such characteristics is available and cannot be easily "adjusted" by taxpayers once their role in determining estimated revenues is known. In the Italian experience, there is some evidence that "independent variables" are subject to manipulation (Pisani, 2004).

This does not disqualify BSA, as far as the scope for manipulation is smaller for the selected "independent" variables than for traditional accounting variables. However, the selection of variables on which to base the estimate is clearly crucial for the successful implementation of the methodology. An associated issue is that BSA determines revenues not income. Therefore, even if revenues are less subject to underreporting, there may be an incentive to inflate costs. This is mitigated by the fact that some cost lines (for instance "cost of goods sold") act as an independent variable in the estimation process. If costs are inflated, this will be partly reflected in an increase in estimated revenues, thereby reducing the incentive to do so.

Establishing the relationship between the independent variables and expected revenues is quite complex and requires some time- and resource- consuming preparatory work. The Italian legislation was introduced in 1993 with the aim of BSA becoming operative in 1995. However, this happened only in 1998. Therefore, establishing some sort of BSA should not be considered as a quick fix, but rather as a medium term process. Due to the complexity of the methodology it is advisable to start on a small scale, applying BSA only to some selected sectors, possibly extending to others once the necessary know-how has been accumulated. A positive aspect of the effort made to formalize in great details the application of BSA is that it reduces the discretion by tax officers, thereby reducing the scope for corruption.

Introducing a BSA does not reduce compliance costs for the taxpayer, quite the opposite. Additional information need to be provided and new forms to be filled in. A positive aspect is that the fact that the presumption is rebuttable provides an incentive to keep proper documentation and accounts, to be able to prove lower actual revenues to the tax authority, thereby improving the quality of information available for traditional tax assessment. The active participation of trade associations at the elaboration stage also increases the informational base on which a BSA is based. It

also increases the perceived legitimacy of this fiscal instrument by the taxpayer.

A definite advantage of BSA is its flexibility as an instrument to fight underreporting. Different sectors and different aspects within a sector can be targeted with specific measures. For instance, in Italy, like in many other countries, undeclared work is particularly common in the construction sector. In 2007 specific consistency criteria were introduced, in an experimental way for the first 2 years, to target this problem. In an agreement with the social partners it was established that total labor cost should account for at least 13.77% of total costs in case of construction of roads and bridges, 22% for renovation of civil buildings and so on for 17 different typologies of construction works. Similar measures are also to be implemented in agriculture and other sectors in which the incidence of undeclared work is high.

4 Minimum social insurance thresholds in Bulgaria

At the end of the '90s, more than one third of labor compensation was estimated to be unreported to fiscal authorities in Bulgaria (Kyle et al., 2001), through unregistered employment relationships and underreporting of earnings. Thus, a part of the workforce was completely underground, while a segment of the workforce that remained in the formal economy minimized payments of social security contributions and other taxes by formally declaring a wage at or near the statutory minimum while receiving additional compensation in cash. In 2002, social security contributions were paid on the basis of the minimum wage for 1.2 millions out of 1.9 millions working on the basis of an employment contract (Neykov, 2003).

To crack down on these activities the Bulgarian government implemented in 2003

two major changes in labor regulation (Neykov, 2003; Pashev, 2006). The first one was the introduction of compulsory registration with the National Social Security Institute of all concluded, amended or terminated employment contracts. In particular, a contract has to be registered before the first day of work. The fact that registration has to take place in advance and not within a certain period after the commencement of the employment relationship makes it easier for labor inspectors to establish a breach of the regulation, as unregistered workers cannot justify themselves by pretending to be on their first day of work.

Another major reform was the introduction of minimum social insurance thresholds (MSIT thereafter) varying according to sector and occupational group. Social security and taxes ought to be paid on the basis of the effective compensation received. Due to underreporting, however, this was not the case for a large part of the workforce, which declared the statutory minimum instead. The reform aimed at increasing the lowest amount on which social security contributions have to be paid and at making it sector- and occupation- specific. Factors like average productivity in a sector or for a specific occupation can be accounted for with differentiated thresholds. Higher compliance for high productivity sectors and occupations can thus be enforced by applying higher minima, without the risk of pricing out workers in lower productivity sectors or occupations.

In the year of their introduction MSIT were differentiated along 9 occupational groups and 48 sectors. The 9 occupational groups are the following:

- legislators, senior officials and managers
- professionals
- technician and associate professionals
- administrative staff
- service workers and sale workers

- skilled agricultural and fishery workers
- craft and related trade workers
- plant and machine operators and assemblers
- elementary occupations

The degree of differentiation along the sectorial dimension has been progressively expanded, with the coverage reaching 68 sectors in 2005 and 73 in 2007. In 2007, therefore, 657 different MSIT could be potentially established. In reality, for many cells in the sector/occupation matrix, the statutory minimum wage (180 leva per month, 126 USD at the average exchange rate for that year) was also used as the minimum social insurance threshold. The highest MSIT for 2007 was fixed at 851 leva per month (596 USD), i.e. approximately 5 times higher than the statutory minimum, for managers in "manufacture of coke, refined petroleum products and nuclear fuel". A similar situation applied in 2009, with 85 categories having a MSIT at the national minimum wage level (240 leva, 171 USD) and the highest threshold, applying to the same category as in 2007, being approximately 5 times higher at 1346 leva per month (957 USD) (Tomev, 2008).

MSIT are negotiated each year with social partners. In case no agreement is reached or no social partner organization exists for a given sector, MSIT can be determined administratively by "expert analysis". In 2005, for instance, an agreement with social partners was reached in 48 out of the 68 sectors for which MSIT were separately determined. In 2008, this figure stood at 45 out of 73 sectors, with an agreed average increase of 26.6% in MSIT applying in 2009 (Tomev, 2008). The importance of social partners is crucial as they are likely to be better informed than the state administration about the effective, as opposed to declared, wages that are prevalent in a sector. This facilitates the fixing of MSIT at the "appropriate level". A too low MSIT compared to effective wages fails to capture most of underreporting.

A too high one, on the other side, endangers competitiveness and employment.

The participation of social partners is likely to be most effective if, as it is the case in Bulgaria, there is a consensus about the need to fight undeclared wages. Trade unions consider the risk that underreporting poses to the social security system, while employers' organizations worry about the unfair competition suffered by compliant businesses (Neykov, 2007). The participation of social partners in the process of determining the thresholds also increases the legitimacy of this fiscal instrument. Moreover, one of the shortcomings of using social insurance thresholds is that they may be regressive if the wage that is actually paid is between the statutory minimum and the higher MSIT. However, in most of the cases in which an agreement among social partners is reached, negotiated minimum social insurance thresholds become sectorial minimum wages for the different occupations through the extension of collective agreements to the whole sector, thus reducing the risk of MSIT above the actual wage.

5 Conclusions and policy recommendations

The policies analyzed in this paper concern the fixing of minimum thresholds to avoid taxpayers declaring an "implausibly" low income. The Bulgarian system is concerned with employed labor and fixes differentiated thresholds for social security contributions according to sector and profession. The Italian system targets small and medium enterprises, self-employed, and professional, by establishing estimated revenues and compensations according to a complex procedure and increasing the probability of being subject to an audit (plus reversing the burden of proof) for those taxpayers declaring lower amounts. A major difference is that in the Italian system the presumption is rebuttable, while this is not the case in the Bulgarian one.

The instruments analyzed in this report are rather flexible and offer the possibility to better target sectors and professions where concealing income from the tax authority is relatively easy, while not introducing too stringent constraints on the rest of the economy. When the capacity of employees to conceal income widely differs across sectors and professions, then establishing differentiated minimum social security thresholds or, quite equivalently, differentiated minimum wages, should be considered. The application of differentiated minima is administratively feasible, even for developing countries (for instance, in Costarica in the period 1988-2000 there were between 19 and 520 minima, according to occupation/skill categories, see Gindling and Terrell, 2007). However, there is the risk of an adverse labor market impact if the level at which the different minima are fixed is not appropriate. This is of course also the case with an undifferentiated statutory minimum wage. If anything, differentiation allows to better calibrate the applicable minimum to the specific conditions of a region or sector or profession, depending on the dimensions along which differentiation will take place.

Differentiated minima may be used to contrast underreporting of wages, part of the so-called "grey economy". Yet, they are ineffective in contrasting completely undeclared labor, participating in the "black economy". The same is true for Business Sector Analysis, as firms and professionals that operate completely outside the legal system are unaffected. However, as it has been seen in the most recent Italian experience in the construction sector, it is possible to use BSA to contrast black work that takes place in registered companies.

An important feature of both the Italian and the Bulgarian experiences is the involvement of social partners. This has the double advantage of improving the informational base on which the policy is founded and boosting its legitimacy among taxpayers. Of course, this requires social partners willing to collaborate with fiscal

authorities in the fight against underreporting. This may not be always the case, but it seems likely that social partners prefer a system in which they have a say to one elaborated and managed solely by the state administration. To fully play their role in the elaboration of measures to fight undeclared work, social partners should be representative and present in all sectors of the economy. Social partners that concentrate their activity in budgetary institutions and big enterprises may not be able to contribute with an in-depth knowledge of business conditions in the parts of the economy in which underreporting is more likely to be widespread, namely among small and medium enterprises. However, the issue should be seen in a dynamic perspective, as the involvement of social partners in the elaboration of fiscal measures that affect their business activity should provide an incentive for small and medium enterprises to participate in employers' organizations, thus gradually improving their representativeness and potential contribution to the fight against undeclared work.

From a tax administration point of view, differentiated social security minima or minimum wages are relatively easy to implement and inexpensive. However, due to the possibility of adverse labor market effects, they should be implemented only after suitable analysis has been conducted. The two crucial variables are the dimensions along which differentiation should take place and the level at which to fix the different minima. Business Sector Analysis is a more complex and time consuming tool that, nevertheless, could prove very effective in fighting underreporting. One advantage of BSA, that makes it potentially applicable also in countries with limited administrative capacity, is that it is scalable, both in terms of the scope of applicability within the economy and in terms of the analytical sophistication of its approach. Thus, BSA could be initially introduced in a limited number of sectors and in an experimental way, so that the necessary know-how can be accumulated and the impact evaluated. If it proves successful, it can then be gradually extended.

As mentioned in the introduction, the enforcement of compliance with tax regulation is a complex task. However, it is not a new task for state administrations around the world and successful international experiences can be fruitfully used as a source of inspiration.

Notes

[1]See Andreoni, Erard and Feinstein (1998), Slemrod and Yitzhaki (2002), and Slemrod (2007) for recent surveys of the literature and Fuest and Riedel (2010) for the case of developing countries.

[2]The assumption is that the tax authority cannot assess and upheld in court a tax liability higher than the true one. To extend the model to situations where this may not be the case, due for instance to ambiguity in the tax code, would be straightforward.

[3]An equivalent narrative is that in an audit, the tax authority may find no evidence at all of tax evasion with probability $H(x)$, which is increasing as the tax liability declared to the authorities increases. Conditional on detection taking place, the density for any given level of income $\hat{y} \in [x, y]$ being discovered is given by $h(\hat{y}) / [1 - H(x)]$.

[4]To simplify the model, I have assumed that audit costs are fixed and hence independent of the taxpayer's gross income. It may be the case that auditing firms or individuals with larger incomes is more expensive. However, even if that is the case, it is unlikely that audit cost increase proportionally with gross income as fixed costs are likely to play an important part in an audit.

[5]See the websites of "Societa' per gli Studi di Settore" (www.sose.it), the Italian Treasury (www.finanze.it) and the Italian tax authority (www.agenziaentrate.gov.it)

for details. This and the following sections are partly based on a report prepared for the World Bank project "Hungary Undeclared Employment".

[6]Unless otherwise indicated, this section refers to Business Sector Analysis in 2007.

References

[1] Andreoni, J., Erard, B., and Feinstein, J. (1998), "Tax compliance", Journal of Economic Literature, vol. 36, pp. 818-860.

[2] Bahl, R., and Bird, R. (2008), "Tax Policy in Developing Countries: Looking Back—and Forward", National Tax Journal, vol. 61(2), pp. 279-301.

[3] Bigio, S., and Zilberman, E. (2010), "Optimal self-employment income tax enforcement", Journal of Public Economics, forthcoming.

[4] Bird, R. (2004), "Administrative Dimensions of Tax Reform", Asia-Pacific Tax Bulletin, vol. 10(3), pp. 134-150.

[5] Convenevole, R., Farina, A., Perinetti, A.R., and Pisani, S. (2006), "Gli effetti dell'applicazione degli studi di settore in termini di ampliamento delle basi imponibili - Relazione per la Corte dei Conti", Documenti di lavoro dell'Ufficio Studi, Agenzia delle entrate, http://www1.agenziaentrate.it/ufficiostudi/documenti/2007.htm

[6] Erard, B. (1997), "Self-selection with measurement errors. A microeconometric analysis of the decision to seek tax assistance and its implications for tax compliance", Journal of Econometrics, vol. 81, pp. 319-356.

[7] Feinstein, J. (1991), "An econometric analysis of income tax evasion and its detection", RAND Journal of Economics, vol. 22, pp. 14-35.

[8] Fuest, C. and Riedel, N. (2009), "Tax evasion, tax avoidance and tax expenditures in developing countries: A review of the literature", Report prepared for the UK Department for International Development (DFID).

[9] Gindling, T.H. and Terrell, K. (2007), "The effects of multiple minimum wages throughout the labor market: The case of Costa Rica", Labour Economics, vol. 14(3), pp. 485-511.

[10] Kyle, S., Warner, A., Dimitrov, L., Krustev, R., Alexandrova, S., Stanchev, K. (2001),"The shadow economy in Bulgaria", Institute for Market Economics, Sofia, http://ime.bg/en/articles/the-shadow-economy-in-bulgaria/

[11] Macho-Stadler, I., and Perez-Castrillo, D. (1997), "Optimal auditing with heterogeneous income sources", International Economic Review, vol. 38(4), pp. 951-968.

[12] Neykov, I. (2003), "New legislation on minimum social insurance thresholds and registration of employment contracts", EIRO, http://www.eurofound.europa.eu/eiro/2003/07/feature/bg0307101f.htm

[13] Neykov, I. (2007), "Employers denounce unfair competition due to grey economy", EIRO, http://www.eurofound.europa.eu/eiro/2007/01/articles/bg0701029i.htm

[14] Pashev, K. V. (2006), "Presumptive Taxation: Lessons from Bulgaria", Post-Communist Economies, vol. 18(4), pp. 399 – 418..

[15] Pisani, S. (2004), "Il triathlon degli studi di settore", Documenti di lavoro dell'Ufficio Studi, Agenzia delle entrate, http://www1.agenziaentrate.it/ufficiostudi/documenti/2004.htm

[16] Reingaum, J., and Wilde, L. (1985), "Income tax compliance in a principal-agent framework", Journal of Public Economics, vol. 26(1), pp. 1-18.

[17] Sanchez, I., and Sobel, J. (1993), "Hierarchical design and enforcement of income tax policies", Journal of Public Economics, vol.50(3), pp. 345-369.

[18] Santoro, A., (2006), "Evasione e studi di settore: quali risultati? quali prospettive?", in: M.C Guerra, A. Zanardi (eds.), Rapporto di finanza pubblica, Bologna: Il Mulino, pp. 297-320.

[19] Scotchmer, S. (1987), "Audit classes and tax enforcement policy", The American Economic Review, vol. 77(2), pp. 229-233.

[20] Slemrod, J. (2007), "Cheating Ourselves: The Economics of Tax Evasion", Journal of Economic Perspectives, vol. 21(1), pp. 25–48.

[21] Slemrod, J., and Yitzhaki, S. (2002), "Tax avoidance, evasion, and administration", in: Auerbach, A., Feldstein, M. (eds.) Handbook of Public Economics, vol. III, Elsevier, Amsterdam.

[22] Thuronyi, V. (1996), "Presumptive Taxation", in: Thuronyi, V. (ed.), Tax Law Design and Drafting (volume 1), International Monetary Fund, Washington, D.C.

[23] Tonin, M. 2007. Essays on Labour Market Structure and Policies, Institute for International Economic Studies, Monograph Series No 59, Stockholm.

[24] Yitzhaki, S. (1974),"A Note on 'Income Tax Evasion: A Theoretical Analysis'", Journal of Public Economics, vol. 3 (May), pp. 201–2.