sciendo

EUROPEAN SPATIAL RESEARCH AND POLICY

Volume 25 2018 Number 1

http://dx.doi.org/10.18778/1231-1952.25.1.02

ARTICLES

Alfonso CARFORA[*], Rosaria Vega PANSINI[*], Stefano PISANI[*]

SPATIAL DYNAMIC MODELLING OF TAX GAP: THE CASE OF ITALY

Abstract. This paper analyses the determinants of regional tax gap in Italy testing if tax evasion is characterised by spatial persistence. The size of spatial correlation in regional tax gaps has been tested and the role of additional determinants of evasion over the period 2001–2011 has been estimated. Using a dynamic spatial panel model, it is shown that regional tax gap is determined by tax evasion in neighbouring regions and is characterised by spatial persistence. Results make it possible to draw a taxonomy of the determinants of regional tax gap: contextual factors and operational factors linked to the relative efficacy of tax evasion contrasting policies and geography.

Keywords: determinants of tax gap, spatial econometrics, panel estimation.

1. INTRODUCTION

The success of an auditing scheme and a tax payment enforcement policy by tax revenue agencies depends on a wide range of factors, like for example the accountability and the integrity of civil servants, the feasibility of the tax rationale and the effectiveness of the fiscal administration. Among these, the capacity to measure tax evasion and identify its determinants are surely of great importance to monitor the progress and to positively impact on tax compliance. Can geography and proximity represent an additional factor influencing tax evasion? Across regions of the same country, can the level of tax evasion be affected by the level of evasion of neighbouring regions?

[*] Alfonso CARFORA, Rosaria Vega PANSINI, Stefano PISANI, Italian Revenue Agency, Via Cristoforo Colombo, 426, 00145 Rome, Italy; e-mails: alfonso.carfora01@agenziaentrate.it, rosariavega.pansini@agenziaentrate.it, stefano.pisani@agenziaentrate.it

The determinants of tax evasion have been extensively analysed by previous and recent literature (Alm, 2012; Yitzhaky, 1974; Clotfelter, 1983; Marino and Zizza, 2012). Among factors of non-compliance, geography has been also added to the picture. It has been considered either to control for regional fixed effects (Richardson, 2006; Schneider and Enste, 2000) with respect to tax collection efficiency or tax settings (Depalo and Messina, 2011) or in evaluation of the shadow economy (Williams and Windebank, 2011; Schneider and Williams, 2013). Finally, spatial adjustments have been included into a MIMIC approach to analyse regional differences in shadow economy in Europe (Herwartz *et al.*, 2015).

The aim of this paper is to explicitly address the role of geography and proximity in modelling regional levels of tax gap for Italy. Our analysis is new in different respects. First, the attention is on tax gap rather than on shadow economy, since the tax gap is the most relevant aggregate of shadow economy. Second, the analysis is focused on Italy thank to the availability of an eleven-year panel of regional tax gap estimates that allows us to explore dynamic spatial correlations and proximity issues in fiscal non-compliance. Third, the empirical analysis is based on the estimation of a spatial dynamic panel model including the share of evaded taxes of neighbouring regions as additional explanatory variable.

As for the measure of non-compliance, we use the share of tax gap on potential tax revenues that are the sum of taxes paid and tax gap. Our dependent variable can be considered as a measure of tax gap propensity, i.e. how much each taxpayer evades for each unit of tax liability. We then estimate the size of spatial correlations among Italian regional tax gap propensities over the period 2001–2011, and three model specifications starting from a pooled OLS to a dynamic panel showing that estimated coefficients may be inefficient in presence of spatial correlation of residuals.

The rest of the paper is organised as follows. Section 2 sketches some evidence on the geographical distribution of tax gap propensities among Italian regions. Section 3 provides some theoretical considerations on the role of geographical aspects in explaining tax gaps. Section 4 describes the empirical strategy while section 5 presents the data. Section 6 contains the results. Finally, Section 7 concludes providing few suggestions for further research and discusses the policy implications of the results.

2. GEOGRAPHY AND THE TAX GAP: SOME EVIDENCE FROM ITALIAN REGIONS

Tax gap is used in this study as a measure of tax evasion. Tax gap is defined as the difference between the potential tax yield that could be collected if no taxpayers would voluntary cheat tax payment and the actual tax revenues. We use a panel database of regional tax gaps estimated by the Italian Revenue Agency (hereafter, IRA)[1] from 2001 to 2011. The estimation follows the 'top-down' approach based on the comparison of the declared income tax base with data from National Accounts on value added, which represents the potential tax base.[2] The tax gap (E) is then calculated as the ratio of monetary level of total tax gap[3] and the potential amount of tax revenues (PTR), i.e. the sum of taxes actually paid and the tax gap itself:

$$E_{it} = \frac{Tax\ Gap_{it}}{PTR_{it}} \qquad [2.1]$$

Where *i* indicates the region and *t* the tax year. The ratio in [2.1] can be seen as an indicator of the regional propensity of non-compliance and a proxy of tax evasion as it measures the amount of each monetary unit of tax gap per each unit of potential tax revenues:

$$E_{it} = \begin{cases} = 0 & Absence\ of\ \ Evasion \\ 0 < E \leq 1 & Tax\ Gap \leq PTR \end{cases} \qquad [2.2]$$

Figure 1 shows the 2001–2011 distribution of regional tax gaps.[4] The analysis of regional box plots representing the time-variation of tax gap shows the heterogeneity of the phenomenon among the Italian regions among the two extremes of Lazio (0.13)[5] and Basilicata (0.47). Moreover, in some regions, like Emilia Romagna, Lazio, Liguria and Tuscany, the tax gap remained quite stable while others, like Calabria, Sardinia and Sicily experienced a large variation in TG values. The TG's time variation can be due to large yearly differences either of potential tax revenues or of tax gap. It is also interesting to note that southern regions[6] experienced a large dispersion of tax gap propensities index trough time. Figure 1 shows also very few outliers represented by some yearly observations

---

[1] The 20 Italian regions are defined by NUTS 2 classification level and have administrative power.

[2] For a detailed description of the methodological issues of tax gap measurement, see Braiotta *et al.*, 2015 and D'Agosto *et al.*, 2014.

[3] By total tax gap we refer to the value of the evasion estimated on taxes under the duty of the IRA, i.e. the sum of VAT, personal income tax (namely, IRPEF), corporate income tax (namely, IRES) and tax on production activities (namely, IRAP).

[4] Tax gap and potential tax revenues are calculated by the Italian Revenue Agency depending on firms' registered address. The geographical distribution of these two measures can be different if calculated using the region where firms' plants are located.

[5] It should be noted that Lazio includes Rome and therefore large share of its potential tax revenues of are represented by public administration and central government. This sector of economic activity has a tax gap equal to 0.

[6] Basilicata, Calabria, Campania, Puglia, Sardegna and Sicily.

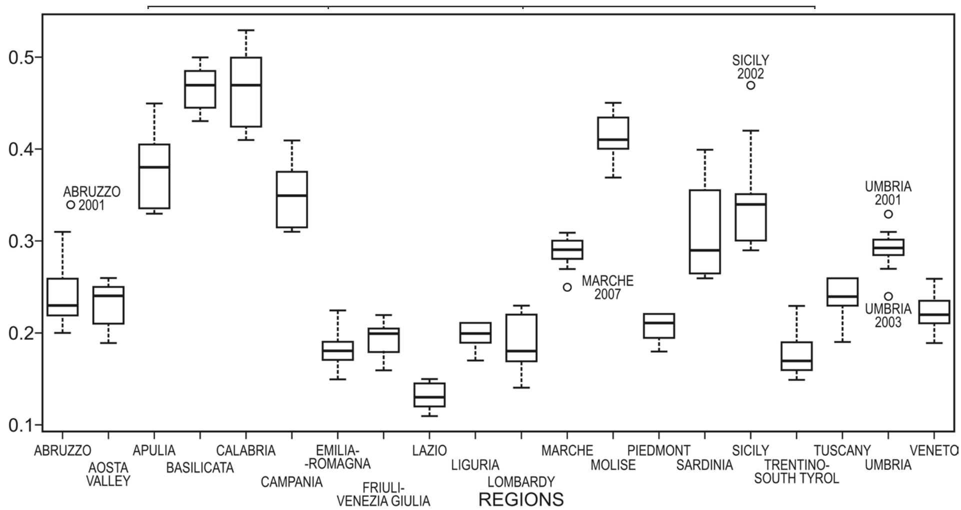

Fig. 1. 2001–2011 distribution of Italian regional Tax Gaps

Source: authors' calculation based on IRA estimates

of TG in four regions. Evidence of regional differences in tax gap shares appears also by comparing the 2001–2011 descriptive statistics,[7] as shown in Tab. 1 and Fig. 2. Southern regions experience the highest incidence of tax gap on potential tax revenues while Lazio, Emilia Romagna, Lombardy and Friuli-Venezia Giulia are those with the lowest average propensities. Between these two extremes, a graduation can be observed mainly among regions in the Center Italy. Even in the block of southern regions, a small variation in TG share can be observed between Campania, Apulia and Sicily and Molise, Calabria and Basilicata. Descriptive analysis provides a first evidence of the relevance of geography and proximity in explaining tax gap shares. Our empirical analysis is meant to analyse if neighbourhood effects can be significant in explaining a regional share of tax gap with respect to potential fiscal revenues. Such descriptive geographical distributions suggest that a 'region-specific' effect and spatial correlations should be considered in the econometric model.

[7] For each *i*-region we calculate the ratio between the average *tax gap* (period 2001–2011) and the average *PTR* (period 2001–2011) as $\bar{E}_1 = \frac{E(TaxGap_i)_{01-11}}{E(PTR_i)_{01-11}}$

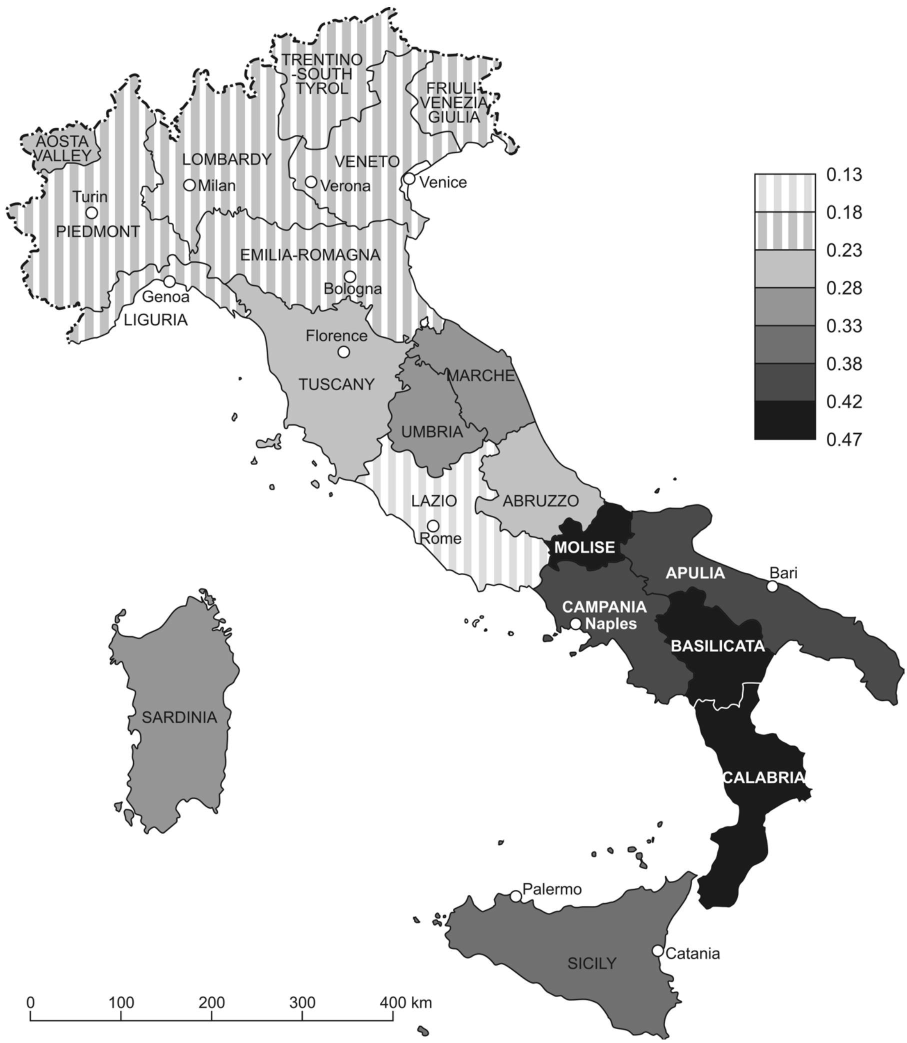

Fig. 2. Regional distribution of average 2001–2011 tax gaps

Source: authors' calculation based on IRA estimates

Table 1. Regional Tax Gaps (E): Descriptive statistics

Region	*mean*	*sd*	*IQR*	*median*
Abruzzo	0.331	0.079	0.068	0.293
Aosta Valley	0.299	0.040	0.064	0.319
Basilicata	0.871	0.081	0.144	0.875
Calabria	0.878	0.149	0.243	0.871
Campania	0.546	0.093	0.132	0.527
Emilia-Romagna	0.217	0.024	0.034	0.221
Friuli Venezia Giulia	0.242	0.032	0.035	0.251
Lazio	0.151	0.018	0.032	0.151
Liguria	0.246	0.019	0.024	0.256
Lombardy	0.235	0.050	0.078	0.215
Marche	0.400	0.034	0.042	0.396
Molise	0.704	0.070	0.090	0.686
Piedmont	0.259	0.026	0.045	0.269
Apulia	0.604	0.113	0.180	0.609
Sardinia	0.460	0.120	0.197	0.399
Sicily	0.531	0.152	0.121	0.514
Tuscany	0.317	0.033	0.040	0.322
Trentino-South Tyrol	0.218	0.037	0.048	0.200
Umbria	0.412	0.046	0.042	0.425
Veneto	0.286	0.033	0.040	0.288

Source: authors' calculation based on IRA estimates.

3. THE MODEL

The reference theoretical framework is the well-known model of Allingham and Sadmo (1972), recently extended by Alm and Yunus (2009) and Di Caro and Nicotra (2014) to consider also spatial aspects of tax evasion. It considers tax evasion as a result of an individual maximization choice in which the taxpayer decides how much to evade based on the level of his expected utility. As observed by Andreoni *et al.* (1998), this general model is not always able to predict a growing empirical evidence on the role of other factors, like moral sentiments, guilt and shame as well as contextual characteristics. An extension to the model is then proposed in order to comprise other variables related to socio-economic and contextual factors, and variables assessing the intervention of the fiscal administration into the general theoretical framework.

Let y_{it} be the individual income, which is unknown to the tax authority. It should be underlined that average taxpayer decides where to locate production and consumption given business opportunities provided by a region's economic system. Such decision is influenced by a set of exogenous economic and institutional characteristics that will be included in the empirical analysis. Let τ_{it} be the constant tax rate on personal income and p_{it} the probability of being audited. If the taxpayer is caught cheating she/he is charged a penalty rate θ_{it} on the evaded income E_{it}. We think that the penalty rate is directly linked to the characteristics of the production system and the efficacy of the auditing process by the fiscal administration. Moreover, the average audited taxpayer faces an additional cost that tax morale and other sentiments (like guilt and shame) represent in individual choice to evade. We indicate this extra cost with δ_{it}. The average taxpayer *i* when deciding how much income to evade takes into account also the amount E_{jt} evaded by another average taxpayer *j* (with $j \neq i$). Individual after tax income in the state 'not caught cheating' is:

$$Y_{i_1}^{AT} = y_{it} - \tau_{it}(y_{it} - E_{it})$$

while, in the state 'caught cheating', it is:

$$Y_{i_2}^{AT} = y_{it} - \tau_{it} y_{it} - \theta_{it}\tau_{it}E_{it} - \delta_{it}E_{it}$$

Each taxpayer is characterised by a von Neumann-Morgestern (1953) expected utility:

$$(U_{it}) = (1 - p_{it}(E_{jt}))U_{it}(y_{it} - \tau_{it}(y_{it} - E_{it})) + p_{it}(E_{jt})U_{it}(y_{it} - \tau_{it}y_{it} - \theta_{it}\tau_{it}E_{it} - \delta_{it}E_{it})$$

where $p_{it}(E_{it})$ is the probability of being caught cheating conditional on the evasion of the average taxpayer *j*. The optimal level of tax evasion E_{it} is obtained from the first order conditions of the maximization problem:

$$(1 - p_{it}(E_{jt}))U'_{it}(Y_{i_1}^{AT}) + p_{it}(E_{jt})U'_{it}(Y_{i_2}^{AT})$$

The solution to the taxpayer's utility maximization can be written in the general functional form:

$$E_{it} = f(E_{jt}, X_{it}, C_{it}) \qquad [3.1]$$

where X_{it} represents the set of variables related to the IRA enforcement that influence the individual's choice to evade and C_{it} the contextual factors describing the socio-economic characteristics of the environment with which the taxpayers interacts. This set of exogenous variables refers to those socio-economic and institutional characteristics, to the indicators of the efficacy of auditing policy and to moral sentiments that may drive the individual's decision to evade taxes.

In this paper, the focus is on regional tax gap that depends on tax gap of the neighbour regions and a set of covariates. In the theoretical model we then refer to the average taxpayer *i* located in region *i*. Moreover, as recalled in section 2, tax evasion is measured as tax gap share of potential tax revenues. Thus, in the section devoted to the empirical analysis, tax gap for region *i* in time *t* will be indicated by E_{it}.

4. ECONOMETRIC SPECIFICATIONS

Studying spatial correlations raises several econometric issues. First, if proximity affects tax gap and firms react to evading decisions of neighbour regions, then the choice to evade taxes is endogenous and correlated with the residual term (u). In fact, there can be unobserved characteristics like institutional environment and tax morale that can be spatially correlated among bordering regions. Second, if taxpayers in neighbouring regions are subject to correlated random shocks, regional tax gaps could be correlated. If we omit the spatial dimension, the outcome of a model estimating the determinants of tax gap propensities could be the presence of spatial dependence in the residuals.

Therefore, tax gap shares is modelled in a dynamic spatial panel framework. Moreover, different non-spatial and spatial model specifications are estimated to test the hypothesis of spatially correlated tax gaps,

4.1. Specification 1: Pooled OLS

A preliminary pooled linear OLS analysis is performed following the specification:

$$E_{it} = \alpha + x'_{it}\beta + u_{it} \qquad [4.1]$$

where u_{it} is a random disturbance term of mean 0. This specification does not take into account the spatial error dependence. Regional tax evasion depends only on its own regional characteristics. Model [4.1] may provide a misleading evidence on determinants of tax gap. Moreover, from an econometric point of view, it can be mispecified and lead to inconsistent estimated coefficients.

4.2. Specification 2: Static and dynamic panel model

The presence of heteroskedasticity in the residual component of the pooled OLS model and issues related to the spatial features of the dependent variable lead us to modify [4.1] with a specification that explicitly considers an individual (region-specific) unobserved effect in the residual component:

$$E_{it} = \alpha + x'_{it}\beta + u_i + \varepsilon_{it} \qquad [4.2]$$

The idiosyncratic error ε_{it} is supposed to be independent with regressors x_{it}, while the individual (time-invariant regional) error component u_{it} may be correlated with regressors. In this case, equation [4.2] becomes:

$$E_{it} = \alpha + x'_{it}\beta + \varepsilon_{it} \qquad [4.3]$$

and consistent estimates can be obtain with the estimation of a fixed effects model. If the error component is uncorrelated with explanatory variables, a random effects model should be used to estimate the [4.2]. This is estimated using the family of generalised least squares (GLS) estimators in order to avoid correlations across composite error terms (within individuals). We use a IV approach such as the GMM estimator by Arellano and Bond (1991) to control for serial correlation of disturbances and when the time-lagged dependent variable is included among regressors:

$$E_{it} = \alpha + \lambda E_{it-1} + x'_{it}\beta + u_i + \varepsilon_{it} \qquad [4.4]$$

4.3. Specification 3: Spatial panel model

Even though model [4.4] considers the time persistence, it still neglects spatial correlation in the residual component. In order to examine the effect of proximity, i.e. time and spatial correlation, a spatial dynamic panel model should be specified. In the literature on spatial statistics (Anselin, 1988), the extent of cross-section correlation is measured with respect to a given 'spatial matrix' W that is a nonnegative $N \times N$ matrix (where N is the number of regions) of known constants describing the spatial arrangement of the units in the sample. The non-zero elements of the matrix indicate whether two locations can be considered neighbours. As a consequence, the element w_{ij} indicates the intensity of the relationship between cross sectional units i and j. By convention, the diagonal elements w_{ii} are all set to zero to exclude self-neighbours. This weighting spatial matrix is not symmetric and is generally used in a row standardised form.

The weights of the Italian regional spatial matrix (W) can be obtained calculating proximities using different algorithms. Among these, the most commonly used are: the k-nearest neighbours algorithm, the contiguity weights matrix ($w_{ij} = 1$ if regions i and j have a common boundary; otherwise $w_{ij} = 0$) and the distance-based binary weights matrix ($w_{ij} = 1$ if the distance between regions i and j is less than a threshold cut-off distance, otherwise $w_{ij} = 0$). We discarded the second due to the presence of islands (without common boundaries with other regions) in the sample, and the third since the setting of the thresholds is highly arbitrary. According to the geo-spatial coordinates,[8] the k-nearest neighbours algorithm considers that $w_{ij} = 1$ if the geo-

[8] Shape files with georeferenced data of Italian data are available at this link: http://www.gadm.org/.

graphical center of region *j* is one of the *k* nearest[9] to region *i*, otherwise $w_{ii} = 0$. The main weakness of this method is represented by the choice of the optimal *k*-number of neighbours. In this analysis, we set the choice of the parameter *k* of *W* matrix following a stepwise procedure. In the first step, after setting a range of possible values for *k* from 1 to 10, ten *W* matrices are constructed, one for each value of *k*. In the second step, using an iterative procedure, a Moran's *I*-index is calculated for each *W* matrix. In the final step, the optimal choice is the value of *k* that gives a Moran's index closest to the average value of those calculated in the second step. Such value of *k* is then used in the construction of the *W* matrix employed in the estimation. The w_{ij} elements of the spatial weighting matrix *W* are used to measure spatial correlations. Matrix *W* is employed to calculate the Global Moran *I*-index (Moran, 1950):

$$I = \left(\frac{N}{\sum_i \sum_j w_{ij}} \right) \left(\frac{N \sum_i \sum_j w_{ij} (E_i - \bar{E})(E_j - \bar{E})}{\sum_i (E_i - \bar{E})^2} \right) \qquad [4.5]$$

that measures the degree of proximity, i.e. if high values of a characteristic are located near other high values and vice-versa. It varies from -1 to 1 and its expected value equals $-1/(N-1)$ under the null hypothesis of no spatial autocorrelation. The nearer the *I*-index to the value of 1, the stronger positive spatial autocorrelation is, while high negative values signal a strong negative spatial autocorrelation. The presence of spatial dependence in the residuals of random and dynamic random models makes GLS and GMM estimators inefficient. In order to overcome this risk, we use the SARSAR specification [4.6] originally suggested by Kelejian and Prucha (1999) to capture spatial interactions across units and over time. In structural form, the SARSAR model can be written as:

$$E_{it} = \alpha + \lambda \sum_{j \neq i}^{N} w_{ij} E_{jt} + \rho \sum_{j \neq i}^{N} w_{ij} u_i + \varepsilon_{it} \qquad [4.6]$$

We extend the specification used by Kapoor *et al.* (2007) including both the spatial lagged dependent variable and spatial error components. We then estimate model [4.6] (Mutl and Pfaermayr, 2011) using a two-step procedure. First, a within and a between two stage least squares coefficients are estimated. The two sets of corresponding residuals are then used in the spatial generalised moments estimator (GM) where the moments conditions are modified accordingly.

[9] The choice of the parameter *k* of *W* matrix follows a stepwise procedure. After setting a range of possible values for *k* from 1 to 10, in the first step we construct ten *W* matrices, one for each value of *k*. In the second step, using an iterative procedure we calculate a Moran's *I*-index for each *W* matrix. In the final step, the optimal choice is the value of k that gives a Moran's index closest to the average value of those calculated in the second step. Such value of *k* is then used in the construction of the *W* matrix employed in the estimation.

5. DATA

The Italian Revenue Agency (IRA) produces annual estimates of regional tax gaps and has a large database of indicators for internal use in taxpayers' risk analysis. We select two additional indicators from the IRA's database that are directly connected to the objective of the IRA to increase tax compliance: *IRA_ss* and *IRA_enforcement*. The first is the ratio of the number of not congruous taxpayers on total taxpayers who joined the 'studi di settore'.[10] This is an auditing scheme employed to determine the amount of revenues and taxes that self-employed and small firms should pay.[11] The adherence to this scheme, which in Italy is voluntary and not mandatory, implies the determination of the number of taxpayers that are congruous with its criteria.[12] We use the one-year lag of *IRA_ss* because the effects of the application of this scheme are perceived by taxpayers during the following fiscal years. Not congruous taxpayers have a higher probability to be audited by the IRA. We expect a negative sign for the estimated coefficient as a higher number of not congruous taxpayers directly relates to the strength of the audit scheme and its capacity to detect tax gap.

The variable *IRA_enforcement* is the ratio between the number of audited taxpayers and total taxpayers. It can be considered as a proxy of the probability of a generic taxpayer to be audited and be used as a measure of IRA's enforcement. We expect a negative coefficient: a higher probability to be audited should lower tax gap propensity.

Variables *Q_Agriculture* and *Q_Industry* indicate the value added from agriculture and manufacturing sectors on regional total value added. *Q_Self-employed* is the regional proportion of self-employed on total employees. These variables have been inserted to describe the feature of the regional production sector. We expect a positive sign for agriculture and self-employment quotas and a negative sign for industry given the different sectoral incidence of tax gap on potential tax revenues. We insert the average number of employees by firm as a proxy of firm size (*Size*). As the Italian production system is characterised by a high number of very small enterprises, we expect a negative sign for this variable. In fact, the bigger the firm size the less the opportunities to hide part of the value added produced.

[10] Endogeneity issues are eluded as this variable measures the ratio between not congruous taxpayers and total taxpayers subjected to 'studi di settore', it does not refer to any amount of collectable and evaded taxes.

[11] For details on the effects of 'studi di settore' audit scheme on tax compliance, see Santoro and Fiorio (2011).

[12] Since their institution in 1993 by law n. 427, 'studi di settore' require that taxpayers subject to this audit scheme must attach to their tax return file a form containing information required in order to estimate their revenues. While they are obliged to fill in the form, they are not obliged to respect the criteria of congruity and coherence.

The variable *Crime_pc* is the number of crimes (weighted by their seriousness assigned by the Italian penal law) committed by organised crime associations per every 1000 inhabitants. It has been included as a measure of social riskiness of the region and, in general, as a dimension of social capital. In fact, the variable does not exclusively include economic crimes but also, for example, murders, robberies and extortions. We expect a positive sign of the coefficient because the higher crime rate of the region the more the opportunities to conceal production activities and the higher the tax gap propensity. *D_Gdp* is an indicator of the time-variation of regional GDP. It is a binary variables taking value 1 if in the *t*-year an increase of regional GDP[13] occurred, 0 otherwise. It has been included as a proxy of the business cycle and employed in place of the actual regional GDP growth rate to avoid multicollinearity issues. We do not have a priori for this variable as its sign indicates whether tax gap propensity is pro- or counter-cyclical.

Pos-ATM_pc is the log of the total number of points of sale terminals and ATM machines per resident. It can be considered as a proxy of the diffusion of the electronic money in the area. The use of *Pos_ATM* as a determinant of tax gap is justified by the fact that it can be connected with the use of cash in undeclared business operations. In fact, we can expect that the higher tax gap especially in sectors connected to the use of cash (like for example, retail), the lower the transactions with electronic money. We expect a negative coefficient as a more spread use of trackable means of payment electronic money should reduce tax gap. *Deposits_pc* is the log of per capita amount of bank deposits with banks located in the region. If we think of per capita amount of bank deposits as an indicator of financial wealth, we can expect higher tax evasion the higher financial wealth. For these reasons, we can question about possible endogeneity issue connected to the use of this variable that could include tax gap. Nevertheless, the use of a panel approach allows us to avoid endogeneity problems. Finally, bank deposits have been inserted as they are often used by revenue agencies as one possible indicator to select taxpayers to be audited.

Tax_amnesties measures the yearly amount of taxes paid by taxpayers after being qualified for the tax amnesty and shelter.[14] It has been inserted in logs due to large disparities in regional values.[15] The effect of tax amnesties on tax evasion is controversial (Alm *et. al.*, 1990). It can be negative if, for example, additional resources are invested in increasing audits and enforcement after the tax write-off. It can be positive , especially in the long run, if they are perceived as an incentive to evade by taxpayers. We expect a positive sign given the hypothesis that regions with a higher level of taxes paid after a write-off indicate a higher incidence of tax gap.

[13] At constant price, base year 2005.

[14] The two main tax debt write-offs has been approved by law n. 289 of December 2002 (fiscal amnesty) and law n. 102 of August 2009 (tax shelter).

[15] We use the log of total amount instead of per capita values as there are regions and years in which tax paid for amnesties is null.

6. RESULTS

Results of the estimation of different model specifications, diagnostics and correlation tests are shown in Tab. 2. Given that estimated results are similar in terms of sign of the coefficients, their effect is only interpreted for the final specification.

Table 2. Tax gap determinants

Variable	OLS	Random	Dynamic	SARSAR
Intercept	-0.002	0.088		0.077*
	(0.074)	(0.090)		(0.034)
lag(*TG*)			0.110	0.553***
			(0.084)	(0.062)
lag(*IRA_ss*)	-0.108*	-0.103***	-0.120***	-0.052***
	(0.044)	(0.027)	(0.024)	(0.015)
IRA_enforcement	-0.119	-0.287**	-0.278	-0.233*
	(0.172)	(0.100)	(0.162)	(0.093)
Q_Agriculture	1.721***	0.733	0.408	0.795**
	(0.299)	(0.671)	(0.789)	(0.297)
Q_Industry	0.155*	-0.269*	-0.471**	0.005
	(0.063)	(0.108)	(0.170)	(0.068)
Q_Self-employed	0.389***	0.623**	0.308	0.54***
	(0.112)	(0.221)	(0.280)	(0.098)
Size	-0.062***	-0.058***	-0.024	-0.046***
	(0.006)	(0.010)	(0.013)	(0.005)
POS-ATM_pc(log)	-0.092***	-0.091***	-0.095**	-0.039***
	(0.013)	(0.017)	(0.030)	(0.010)
Deposits_pc(log)	-0.003	0.007	0.042	0.015
	(0.020)	(0.025)	(0.030)	(0.011)
Crime	0.000	0.001**	0.000	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
D_gdp	-0.015*	-0.015***	-0.014***	-0.011***
	(0.006)	(0.003)	(0.003)	(0.002)
Tax_Amnesties(log)	0.001	0.004***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)

Notes: Robust standard errors are in parentheses; Significance level: *** 0.1%, ** 1%, * 5%.

Table 2. cont.

Variable	OLS	Random	Dynamic	SARSAR
Diagnostics				
Adj.R2	0.867	0.583		
θ		0.824		0.683
ρ				-0.688
Sargan (p-value)			0.999	
Cross sectional correlation				
Pesaran CD		4.423	2.801	-0.241
		(0.000)	(0.005)	(0.809)
Local cross sectional correlation				
Pesaran CD (p)		7.101	2.642	0.009
		(0.000)	(0.008)	(0.993)
Baltagi Song, Koh				0.709
				(0.478)
Unobserved effects				
Breusch-Pagan	341.604			
	(0.000)			
Wooldridge	3.013			
	(0.002)			
Hausman		13.581		
		(0.257)		
Serial correlation				
Baltagi and Li		33.190		
		(0.000)		
Breusch–Godfrey		47.075	19.306	6.171
		(0.000)	(0.007)	(0.723)
AR (1)			-3.056	
			(0.001)	
AR (2)			0.945	
			(0.172)	

Note: *p*-values are in parentheses.

Source: authors' calculation based on IRA estimates.

6.1. Specification 1: Possible factors explaining tax gap

Even though OLS coefficients (columns 1) are mostly significant and with expected signs, they ignore the longitudinal features of the dataset. Thus, pooling estimation may suffer from mis-specification. Diagnostics on the residuals and the Breusch-Pagan and Wooldridge tests suggest that OLS estimators are inconsistent, given the correlation between the residual component (u_{it}) and the covariates. The variability of the individual error term with respect to total variability is very large ($\theta = 0.824$). Such evidence suggests the opportunity to use an alternative estimator that takes into consideration the panel dimension of our data.

6.2. Specification 2: Dynamic determinants of tax gap

The coefficients estimated using a random effects model specification are shown in column 2. The results of the Hausman and the Wooldridge tests are also included. They confirm the presence of unobserved individual (regional) effects. The Hausman test does not allow to reject the null hypothesis of equal coefficients between FE and RE models. As in the OLS model, RE model coefficients have the expected sign and are highly significant. In order to check if a static panel model is a right specification, we test the assumptions on serial and cross-sectional independence of disturbances, i.e. if u_{it} is independent with u_{it+k}, for $k = 1,\ldots T$, and if u_{it} is independent with u_{ij} every $i \neq j$. We use the Baltagi-Li (Baltagi and LI, 1995) and Breusch-Goedfrey (Godfrey, 1978) tests for serial correlation and the Pesaran CD test (Pesaran, 2004) for cross-sectional correlation. The Pesaran CD test is implemented on the residuals of model [4.2] that has been pre-treated as an AR structure to avoid serial correlation. Results suggest that the null hypothesis of time and cross-sectional independence cannot be accepted. We then use a dynamic panel specification [4.4] and a GMM estimator à la Arellano and Bond.

Column 3 of Tab. 2 shows the results of the estimation of the [4.4]. The coefficient of the $lag(TG)$ indicates the presence of a significant (as expected) relationship between the tax gap shares at time t and $t+1$ and, consequently, that tax gap is time-persistent. The Sargan test supports our choice of the instruments included in the model, i.e. the lagged values of the dependent variable. Nevertheless, the coefficient of lagged dependent variable is significant only at 10% level. Moreover, given the results of AR(1), AR(2) and of the Breusch-Goedfrey and Pesaran CD tests, specification [4.4] still does not allow to correct serial and cross sectional dependence in the disturbances. These results explain the bad performance of the dynamic specification: due to the presence of sectional dependence the lagged dependent variable is a weak instrument for the different GMM dependent variable (Blundell and Bond, 1998).

6.3. Specification 3: Proximity as an explaining determinant of tax compliance

The presence of serial and cross-sectional correlation raises issues about the consistency of coefficients and suggests switching to a model that explicitly considers the spatial dimension of data. It first is checked if spatial correlation exists using the values of Moran's I-statistics calculated for Italian regional TGs for every year (2001–2011) contained in Tab. 3. Results of the Moran's test (Cliff and Ord, 1981) indicate that the null hypothesis of no spatial correlation between regional TGs cannot be accepted. A local variant of the CD test (Pesaran *et al.*, 2011) is also performed using matrix W to test the null of no cross-sectional dependence[16] (Millo, 2014). Results show that cross sectional correlation of residuals of the random and the dynamic random models can be treated as spatial correlation. Regional tax gap shares are characterized by spatial persistence. As shown in Fig. 1 and 2, regions with high values of average TG are located near regions with equally high values and vice versa.

Table 3. Spatial autocorrelation of regional tax gaps

Year	I-statistics	Moran's test (p-value)
2001	0.642	0.000
2002	0.683	0.000
2003	0.552	0.000
2004	0.510	0.000
2005	0.502	0.000
2006	0.571	0.000
2007	0.560	0.000
2008	0.523	0.000
2009	0.487	0.000
2010	0.442	0.000
2011	0.490	0.000

Source: authors' calculation

The aim of our empirical analysis is then to show that in presence of spatial correlation a SARSAR model should be used in order to obtain consistent estimates (Elhorst, 2003). This choice is also confirmed by tests of cross sectional and local cross sectional dependence, like the Pesaran and Baltagi (Baltagi *et al.*, 2003) tests,

[16] The alternative is the existence of local cross-sectional dependence. Specifically, the dependence between neighbours only is testes on the residuals of models [4.4] and [4.6].

and tests of serial correlation. These were performed on residuals of model [4.6] and results suggest that the null hypothesis of independence and no serial correlation cannot be rejected. Therefore, the analysis of the estimated coefficients by the SARSAR model provides a correct and consistent picture of the role of geography and other covariates in determining regional tax gap propensities.

The coefficient of the spatial lagged dependent variable (*lag*(*TG*)[17]), strongly significant and positive, confirms our a priori idea that tax gap is influenced by spatial persistence too, a result similar to Alm and Yunus (2009) for US: the proximity to a region with high (low) TG share is a significant determinant of the high (low) TG shares of a neighbouring region. The introduction of the spatial lag variable as additional covariate suggests that spillovers effects maybe in action among neighbour regions in influencing taxpayers' attitude toward compliance. The *lag*(*TG*) variable exerts also a relatively high coefficient compare to other covariates. This means that among other factors considered, tax gaps of the neighbour regions is among the strongest in influencing the level of regional tax gap. Moreover, it is also highly significant compared to the results of the dynamic panel model when using a correct model specification in presence of spatial correlation. This result has surely also theoretical implications. The decision to evade is more complex than that described by Allingham and Sadmo model and can be determined also by the fiscal behavior of taxpayers in neighbouring regions.

Variables *IRA_ss* and *IRA_enforcement* are both negative and significant. As expected, the sign of the lagged variable measuring the share of not congruous on total taxpayers who joined the *studi di settore* auditing scheme is negatively correlated with regional tax gap propensities. In fact, not congruous taxpayers have a higher probability to be audited and their share negatively influences the power of the *studi di settore* scheme in improving tax compliance. The coefficient of *IRA_enforcement* reveals that an increase in the probability to be audited reduces regional tax gap propensity. This is also an expected result that indicates a positive effect on compliance rate of IRA activities (Alm, 1999 and Yitzhaki, 1974) and it is in line with results obtained for Italy by Marigliani and Pisani (2014).[18]

Looking at the features of regional productive system, the higher the quota of regional value added produced by the agricultural sector the higher the tax gap propensity. This results are similar to those obtained by the National Statistics Institute (ISTAT) in estimating the underground economy.[19] Irregular employees were about 3 million (14% of total workforce) and they were concentrated mainly in the agricultural sector (over 32%) (Fiorio and D'Amuri, 2006). Moreover, the coefficient

---

[17] For the sake of simplicity of Tab. 2, the variable *lag*(*TG*) indicates alternatively the time-lagged tax gap in the dynamic panel specification and spatial-lagged tax gap in the model.

[18] They use a slightly different indicator for IRA enforcement as they calculate the ratio of the amount of taxes collected by preventing and tackling evasion and total tax gap.

[19] For latest estimates see ISTAT (2010) available at http://www.istat.it/it/archivio/4384.

for *Q_Agriculture* is quite high compared to other variables: a one-percentage increase in the share of value added produced by the agricultural sector increases the average regional tax gap share by almost 0.80. The coefficient of the value added of the industrial sector is not significant in the SARSAR specification and it also changes by sign, level and significance throughout model specifications. We expect a negative sign for this variable, i.e. the higher the value added from industrial sector the lower regional tax gap. This a priori comes from the fact that, with respect to the agricultural sector or self-employed and professional, industrial sector is constituted by firms of medium and big size which are subject to a more complex system of accounting and business registration that can be also easily trackable. Given such features of business conduct in the industrial sector, the coefficient of variable *Q_Industry* can be explained by the fact that, on average, firms belonging to the industrial sector have a lower propensity to evade. In fact, given the Italian fiscal law and tax procedural obligations that bigger firms must meet, opportunities to evade taxes are much lower than for less organised small firms and self-employed professionals. Moreover, ISTAT calculates the value added of the industrial sector by the region where the plant is located that can differ from the region where the firm's headquarter is, which is also the region where tax gap share is imputed to. The coefficient for the variable *Q_Self-employed* is positive and highly significant. The magnitude of the effect on tax gap propensity is the same as *lag*(*TG*) variable: a one percent increase in the share of self-employed professional increases the share of tax gap on average by half percentage point. As in Bordignon and Zanardi (1997), this result reflects one of the peculiarities of the Italian productive system characterised by a large proportion of small firms, professionals and self-employed in the total workforce. Ceteris paribus, tax evading chances are clearly larger for self-employed than for employees (Braiotta *et al.*, 2015).[20] The firm's size (*Size*) has a negative effect on regional tax gap. As stated by the Italian fiscal law, large firms are subjects to several additional duties and obligations in order to complete their tax return forms than smaller firms. Moreover, due to specific characteristics of business conduct[21] and a higher number of controls from IRA, as firm's dimension increases there is less room for tax gap.[22]

---

[20] Such evidence is observed also for UK, where the tax gap from individual in self-assessment is 17% of the tax liabilities and the same percentage for the whole taxpayers is 7% (HMRC, 2015) and Denmark, where the evasion rate for individual with self-reported income is equal to 37% (Klever *et al.*, 2010).

[21] As stated by the Italian law, large firms are subject to a specific 'tutoring' activity by the IRA consisting in a deep and long-lasting fiscal assistance. Moreover, large firms are subject to several additional rules involving the spread use of electronic payment and invoicing in business conduct that limit the opportunities of evasion.

[22] The estimates of tax gap used in this analysis refer strictly to tax evasion connected to the production of goods and services. Hence, they do not consider tax avoidance or frauds that can increase the incidence of tax evasion of larger firms.

The effect of crime is positive and significant even though not as strong as expected. This can be due to measurement errors or to the fact that our variable considers not only economic crimes but also different types of illegal activities. A narrower indicator could have a stronger effect on tax gap, as in the study of Friedman *et al*. (2006). Using data from 69 countries, they show how corruption is associated with more unofficial activity. Nonetheless, the positive and significant coefficient confirms the results obtained by Dell'Anno and Schneider (2006). In fact, if the crime variable indicate either the rate of illegality or the efficacy of the police action, it can also have a negative effect on tax gap, as the cost to participate in illegal activities increases (Eilat and Zinnes, 2000).

The binary variable capturing the time variation of regional GDP is negatively correlated with regional tax gap. Our results show that an expansion of regional GDP is negatively correlated with tax gap. This can be due to the the reduction of demand for underground products as GDP increases and the fact that positive regional GDP growth creates more job opportunities into regular economy.

The effect of the diffusion of electronic money (*Pos-ATM_pc*) on regional tax gap is negative and significant. There is a large literature and extensive evidence on the positive relationship between the use of cash and tax evasion. Demand for cash is also used for estimating the extension of tax evasion (Ardizzi *et al.*, 2014). Therefore, the use of electronic money has a detrimental effect on tax gap.

The amount of bank deposits per resident does not affect tax gap share. If it is considered a proxy of the regional level of wealth, depending on the sign of the coefficient, it can have either a positive (cyclical) or negative (anticyclical) effect on tax gap. In our case, bank deposits are significant in none of model specifications estimated. This can be related to the way in which the variable is constructed as it measures those deposits with banks located in the respective region also of people living in other regions. Since a proportion of these deposits does not contribute to the regional tax gap, the variable is not significant in the SARSAR model.

Finally, the amount of taxes paid after fiscal amnesties (*Tax_amnesties*) has a positive effect on tax gap. The increase in tax paid after a fiscal write-off increases tax evasion, as also in (Marigliani and Pisani, 2014). This result confirms our a priori idea that tax amnesties can implicitly represent an incentive for taxpayers to evade tax payment with the perspective to condone their fiscal obligations in the future.

7. CONCLUDING REMARKS

Despite the wide and extensive literature on the determinants of tax evasion, the effect of geography and proximity has not been yet explored. Moreover, from an empirical point of view, the use of spatial econometrics has not been exploited in the analysis of tax gap. This paper provides an original contribution to the debate on the determinants of Italian tax gap at regional level. Spatial econometrics and a SARSAR specification model have been applied to obtain consistent estimates in presence of spatially correlated residuals. Results support the hypothesis that, on average and ceteris paribus, a region's tax gap propensity is positively determined by the tax gap of neighbour regions. Tax gap is then characterised by proximity and spatial persistence. The results of the empirical analysis enables us to distinguish factors correlated to the level of regional tax gap: contextual factors (the incidence of agriculture on regional value added, the share of self-employed professionals, firm size, crime, electronic money, tax amnesties and GDP growth); operational factors linked to the efficacy of tax evasion contrasting policies (IRA enforcement and renewal of auditing schemes) and proximity factors. We acknowledge that our results need additional research especially in the direction of testing the robustness of estimates at a sub-regional level. Moreover, more detailed information at individual level could be harmful to analyse if mimicking behaviour in tax compliance exists among taxpayers in neighbour regions. Nevertheless, this analysis provides useful suggestions for tax evasion contrasting policies. At the regional level, more cooperation is needed among local offices of the IRA in neighbour regions as well as a deeper coordination among fiscal agencies, the Italian fiscal police and the local government. At the central level, the design of auditing schemes should consider possible cooperation among taxpayers in neighbour regions and proximity as an additional factor in taxpayers risk analysis. Finally, given their spatial features, the tax evasion contrasting policies should also be targeted at production districts and supply chains and take into account employees' migration across regions.

REFERENCES

ALLINGHAM, M. G. and SANDMO, A. (1972), 'Income Tax Evasion: A Theoretical Analysis', *Journal of Public Economics*, 1, pp. 323–338.

ALM, J. (1999), 'Tax Compliance and Administration', *Public Administration and Public Policy*, 72, pp. 741–768.

ALM, J. (2012), 'Measuring, Explaining, and Controlling Tax Evasion: Lessons from Theory, Experiments, and Field Studies', *International Tax and Public Finance*, 19 (1), pp. 54–77.

ALM, J., McKEE, M. and BECK, W. (1990), 'Amazing Grace: Tax Amnesties and Compliance', *National Tax Journal*, 43 (1), pp. 23–37.

ALM, J. and YUNUS, M. (2009), 'Spatially and Persistence in U.S. Individual Income Tax Compliance', *National Tax Journal*, 62 (1), pp. 101–124.

ANDREONI, J., ERARD, B. and FEINSTEIN, J. (1998), 'Tax Compliance', *Journal of Economic Literature*, 36 (2), pp. 818–860.

ANSELIN, L. (1988), 'Spatial Econometrics: Methods and Models', [in:] ANSELIN, L., *Spatial Econometrics Chapter Fourteen*, Dordrecht: Kluwer Academic Publisher, pp. 310–325.

ARELLANO, M. and BOND, S. (1991), 'Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations', *Review of Economic Studies*, 58, pp. 277–297.

BALTAGI, B. and LI, Q. (1995), 'Testing AR (1) against MA (1) Disturbances in an Error Component Model', *Journal of Econometrics*, 68, pp. 133–151.

BALTAGI, B. H., SONG, S. H. and KOH, W. (2003), 'Testing Panel Data Regression Models with Spatial Error Correlation', *Journal of Econometrics*, 117, pp. 123–150.

BLOOMQUIST, K. (2003), 'Tax Evasion, Income Inequality and Opportunity Costs of Compliance', *96th Annual Conference of The National Tax Association*, Chicago, pp. 13–15.

BLUNDELL, R. and BOND, S. (1998), 'Initial Conditions and Moment Restrictions in Dynamic Panel Data Models', *Journal of Econometrics*, 87 (1), pp. 115–143.

BORDIGNON, M. and ZANARDI, A. (1997), 'Tax Evasion In Italy', *Giornale degli economisti e Annali di economia*, 56 (3–4), pp. 169–210.

BRAIOTTA, A., CARFORA, A., PANSINI, R. V. and PISANI, S. (2015), 'Tax Gap and Redistributive Aspects Across Italy', *Argomenti di Discussione of Italian Revenue Agency*, 2, pp. 1–27.

CABALLE, J. and PANADES, J. (2000), 'Tax Evasion and Economic Growth', *Public Finance/ Finances Publiques*, 52, pp. 318–340.

CHIARINI, B., MARZANO, E. and SCHNEIDER, F. (2008), 'Tax Rates and Tax Evasion: An Empirical Analysis of The Structural Aspects and Long-Run Characteristics in Italy', *IZA Discussion Papers*, 3447, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1136252.

CLIFF, A. D. and ORD, J. K. (1981), *Spatial Processes: Models and Applications*, London: Pion Ltd.

CLOTFELTER, C. T. (1983), 'Tax Evasion and Tax Rates: An Analysis of Individual Returns', *Review of Economics and Statistics*, 65 (3), pp. 363–373.

D'AGOSTO, E., MARIGLIANI, M. and PISANI, S. (2014), 'Asymmetries in the Territorial Vat Gap', *Argomenti di Discussione of Italian Revenue Agency*, 2.

DELL'ANNO, R. and SCHNEIDER, F. G. (2006), 'Estimating the Underground Economy by Using Mimic Models: A Response to T. Breusch's Critique', *Economics Working Papers, 2006–07 of Department of Economics University Linz*.

DEPALO, G. and MESSINA, G. (2011), 'Follow the Herd. Spatial Interactions in Tax Setting Behavior of Italian Municipalitie's', *Siecon Working Paper*, http://www.siecon.org/online/wp-content/uploads/2011/04/Depalo-Messina1.pdf.

DI CARO, P. and NICOTRA, G. (2014), *Knowing the Unknown across Regions: Spatial Tax Evasion in Italy*, Social Science Research Network, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2446803

EILAT, Y. and ZINNES, C. (2000), 'The Evolution of the Shadow Economy in Transition Countries: Consequences for Economic Growth and Donor Assistance', *CAER II Discussion Paper No. 83 of Harvard Institute for International Development*, pp. 3–70 .

ELHORST, J. P. (2003), 'Specification and Estimation of Spatial Panel Data Models', *International Regional Science Review*, 26, pp. 244–268.

FIORIO, C.V. and D'AMURI, F. (2006), 'Workers' Tax Evasion in Italy', *Giornale degli economisti e Annali di economia*, 64 (2–3), pp. 247–270.

FRIEDMAN, E., JOHNSON, S., KAUFMANN, D., and ZOIDO-LOBATON, P. (2000), 'Dodging the Grabbing Hand: The Determinants of Unofficial Activity in 69 Countries', *Journal of Public Economics*, 76, pp. 459–493.

GILES, D. E. A. (1999), 'Measuring the Hidden Economy: Implications for Econometric Modeling', *The Economic Journal*, 109 (46), pp. 370–380.

GODFREY, L. (1978), 'Testing against General Autoregressive and Moving Average Error Models when the Regressors Include Lagged Dependent Variables', *Econometrica*, 46, pp. 1293–1302.

HERWARTZ, H., TAFENAU, E. and SCHNEIDER, F. (2015), 'One Share Fits All? Regional Variations in the Extent of the Shadow Economy in Europe', *Regional Studies*, 49, pp. 1575–1587.

ISTAT (2010), 'La misura dell'economia sommersa secondo le statistiche ufficiali. Anni 2000–2008', *Statistics in Brief*, 13, https://www.istat.it/it/files//2011/01/testo integrale20100713.pdf.

KAPOOR, M., KELEJIAN, H. H. and PRUCHA, I. R. (2007), 'Panel Data Model with Spatially Correlated Error Components', *Journal of Econometrics*, 140 (1), pp. 97–130.

KELEJIAN, H. H. and PRUCHA, I. R. (1999), 'A Generalised Moments Estimator for the Autoregressive Parameter in a Spatial Model', *International Economic Review*, 40 (2), pp. 509–533.

KLEVEN, H. J., KNUDSEN, M. B., KREINER, C. T., PEDERSEN, S. and SAEZ, E. (2010), 'Unwilling or Unable to Cheat? Evidence from a Randomised Tax Audit Experiment in Denmark', NBER Working Paper, No. 15769.

MARIGLIANI, M. and PISANI, S. (2014), *Una stima dell'effetto deterrenza esercitato dall'Agenzia delle Entrate*, New York: Italian Revenue Agency, Mimeo.

MARINO, M. R and ZIZZA, R. (2012), 'Personal Income Tax Evasion in Italy: An Estimate by Taxpayer Type', [in:] PICKHARDT, M. and PRINZ, A. (eds.), *Tax Evasion and The Shadow Economy*, Chapter 3, Cheltenham: Edward Elgar Publishing.

Measuring Tax Gaps (2015), HM Revenue & Customs, London, UK.

MILLO, G. (2014), 'Narrow Replication of a Spatio-Temporal Model of House Prices in the USA Using R', *Journal of Applied Econometrics*, 30 (4), pp. 703–704.

MORAN, P. A. P. (1950), 'Notes on Continuous Stochastic Phenomena', *Biometrika*, 1/2, pp. 17–23.

MUTL, J. and PFAFFERMAYR, M. (2011), 'The Hausman Test in a Cliff and Ord Panel Model', *Econometrics Journal*, 14, pp. 48–76.

NEUMANN, J. V. and MORGENSTERN, O. (1953), *Theory of Games and Economic Behavior*, Princeton, NJ: Princeton University Press.

PESARAN, M. (2004), 'General Diagnostic Tests for Cross Section Dependence in Panels', *CESifo Working Paper*, No. 1229.

PESARAN, H. M. and TOSETTI, E. (2011), 'Large Panels with Common Factors and Spatial Correlations', *Journal of Econometrics*, 161 (2), pp. 182–202.

RICHARDSON, G. (2006), 'Determinants of Tax Evasion: A Cross-Country Investigation', *Journal of International Accounting, Auditing and Taxation*, 15 (2), pp. 150–169.

SANTORO, A. and FIORIO, C. V. (2011), 'Taxpayer Behavior when Audit Rules are Known: Evidence from Italy', *Public Finance Review*, 39 (1), pp. 103–123.

SCHNEIDER, F. and ENSTE, D. H. (2000), 'Shadow Economies: Size, Causes, and Consequences', *Journal of Economic Literature*, 38, pp. 77–114.

SCHNEIDER, F. and WILLIAMS, C. C. (2013), 'The Shadow Economy', London: Institute of Economic Affairs.

WILLIAMS, C. C. and WINDEBANK, J. (2011), 'Regional Variations in the Nature of the Shadow Economy: Evidence from a Survey of 27 European Union Member States', [in:] *Handbook on the Shadow Economy*, Chapter 5, Cheltenham: Edward Elgar Publishing.

WOOLDRIDGE, J. M. (2002), *Econometric Analysis of Cross-Section and Panel Data*, Cambridge, MA: MIT Press.

YITZHAKI, S. (1974), 'A Note on Income Tax Evasion: A Theoretical Analysis', *Journal of Public Economics*, 3 (2), pp. 201–202.