Geographical Distribution of Crime in Italian Provinces: A Spatial Econometric Analysis

Maria Francesca Cracolici
and Teodora Erika Uberti

NOTA DI LAVORO 11.2008

FEBRUARY 2008

KTHC – Knowledge, Technology, Human Capital

Maria Francesca Cracolici, *Department of National Accounting and Analysis of Social Processes, University of Palermo and Department of Spatial Economics, Free University*
Teodora Erika Uberti, *Department of International Economics, Institutions and Development (DISEIS), Faculty of Political Science, Catholic University of Milano*

Geographical Distribution of Crime in Italian Provinces: A Spatial Econometric Analysis

Summary

For a long time social sciences scholars from different fields have devoted their attention to identifying the *causes* leading to commit criminal offences and recently lots of studies have included the analysis of spatial effects. Respect to the Italian crime phenomenon some stylized facts exist: high spatial and time variability and presence of "organised crime" (e.g. *Mafia* and *Camorra*) deep-seated in some local territorial areas. Using explanatory spatial data analysis, the paper firstly explores the spatial structure and distribution of four different typologies of crimes (murders, thefts, frauds, and squeezes) in Italian provinces in two years, 1999 and 2003. ESDA allows us to detect some important geographical dimensions and to distinguish crucial macro- and micro-territorial aspects of offences. Further, on the basis of Becker-Ehrlich model, a spatial cross-sectional model including deterrence, economic and socio-demographic variables has been performed to investigate the determinants of Italian crime for 1999 and 2003 and its "neighbouring" effects, measured in terms of *geographical* and *relational* proximity. The empirical results obtained by using different spatial weights matrices highlighted that socioeconomic variables have a relevant impact on crime activities, but their role changes enormously respect to crimes against person (murders) or against property (thefts, frauds and squeezes). It is worthy to notice that severity does not show the expected sign: its significant and positive sign should suggest that inflicting more severe punishments does not always constitute a deterrence to commit crime, but it works on the opposite direction.

Keywords: Crime, Spatial Econometrics

JEL Classification: C21, K42

We are grateful to Miranda Cuffaro, Mario A. Maggioni, Gunther Maier, Jan Oosterhaven, Alberto Quadrio Curzio. Although this is a joint paper, paragraphs 1, 3 and 4 are attributed to Francesca Cracolici while paragraphs 2 and 5 to Teodora Erika Uberti. We acknowledge PREPARE Project ID 029619 for financial support.

Address for correspondence:

Teodora Erika Uberti
Department of International Economics
Università Cattolica del Sacro Cuore
Largo Gemelli 1
Milano 20123
Italy
E-mail: erika.uberti@unicatt.it

1. Introduction

Since long time, the study of the determinants and effects of crime has drawn the attention of scholars from social sciences, notwithstanding, economists turned their attention to illegal behaviour of individuals at the beginning of 1960s.

The starting point of the study of crime in economic terms is the seminal paper by Becker (1968) on “Crime and Punishment” (see also Shaw and McKay 1942; Stigler, 1970; Ehrlich, 1973; Merlo, 2004). Becker’s crime economic model (CEM) is based, namely, on the relation between crime and punishment. According to Becker, a rational individual decides to act illegally on the basis of the costs and benefits connected to it; i.e. the illegal behaviour of individuals could be explained by means of the theory of rational behaviour under uncertainty (Becker 1968).

In 1973, Ehrlich proposed an extension and empirical application of CEM considering a time allocation model. The empirical model proposed by Ehrlich considers the opportunities connected to both punishment and reward; i.e. cost and gains from legal and illegal activities. He verified the relation between the rate of crime and either deterrence and socio-economic variables.

In recent years, the studies on crime have shifted the attention from the strict CEM to *atheoretical* models based on empiricism. Most empirical models include not only deterrence variables, but also socio-economic and demographic determinates that could explain the persistence and high rate of crime like unemployment, age composition of population, increasing income inequality, education, etc. (see Marselli and Vannini 1997; Cezay *et a*l. 1998; Entorf and Spengler 2000; Buonanno and Leonida, 2006; Edmark 2005; Buonanno and Leonida 2006). Further, recently, complex or frontier economic models of crime departing from CEM have been proposed (e.g., see Burdett et al. 2004; Huang et al. 2004; Lochner 2004).

In Italy, the analysis of crime in economic and quantitative terms has received attention only in the last years (see Marselli and Vannini 1997; Marselli and Vannini 2000; Buonanno and Leonida, 2006; Marselli and Vannini 1995). An interesting empirical analysis of economic crime model for Italian regions for the period 1980 to 1989 has been performed by Marselli and Vannini (1997). In order to take into account the phenomenon of criminal organizations characterising the Italian crime, the authors, using a panel approach, investigated the relation among four types of crime (murder, theft, robbery and fraud) and some relevant economic and deterrence variables. They found the Italian crime is characterized by some facts: “(i) the probability of punishment is relatively more effective than both the severity of punishment and the efficiency of police authority in deterring crime; (ii) among the variables representing the opportunity costs of participating in illegal activities, the rate of unemployment, the value of public works started by government, and the proportion of people employed in the service sector have a significant effect; (iii) for three types of crime the regional unobservable component is correlated with the regressors; (iv) spillovers from drug consumption to theft are substantial; (v) with the exception of fraud, the results are in contrast with the predictions of the standard economic model of crime” (pp. 89).

To the exception of the previous studies, it is worthy to notice the analysis of crime in Italy according to an economic prospective has not received a wide attention yet. In other words, whether, on the one hand, scholars have paid attention on territorial disparities of economic development, on the other hand the economic empirical literature has not enough focussed on the differences of crime rates among regions as mirror of the regional differences of socio-economic development (see, e.g., Hsieh and Pugh 1993; Ichiro et al. 1999)..

Following the previous literature, in this paper we analyse and verify the relation between crime and deterrence, economic and demographic variables for the 103 Italian provinces for two years, 1999 and 2003.

The approach we follow here is loosely similar to that taken by Maselli and Vannini (1997). In contrast to the latter, we propose an analysis of crime at a lower territorial scale, i.e. at provincial level, to capture important interregional differences. Our aim is to explore which are the provincial determinants of crime in Italy.

Further, because the spatial analysis of crime has demonstrated that the location of illegal activity can supply relevant insights about the exploration of crime dynamics (see Messner et al. 1999; Anselin et al. 2000), we use spatial econometric tools to assess empirically the crime determinants considering the location of crime. Using a spatial regression model – based on spatial autocorrelation techniques – for cross-sectional Italian data, we investigate if Italian crime is characterized by neighbouring effects, either in terms of geographical contiguity distance based proximity and relational proximities; i.e. if the crime activity spreads through a diffusion process or not.

The paper is structured as follows. In Section 2, by using the methodology of exploratory spatial data analysis (ESDA) identification and interpretation of spatial clustering of different types of crime (murders, thefts, frauds and squeezes) have been made. Section 3 presents some theoretical background and introduces the statistical model and the data used in our empirical application. In Section 4, the empirical findings are presented and interpreted. And, finally, Section 5 concludes the analysis.

2. The Italian crime: an ESDA investigation

As stressed in Marselli and Vannini (1997), Italian crime has particular qualitative and quantitative features: firstly crime activities vary across time and space; secondly "organised crime", like *Mafia*, *Camorra*, *'Ndrangheta* and *Sacra Corona Unita*, has territorial roots in some southern regions. Bearing in mind these important features of the Italian crime phenomenon, we tried to capture the differences across time, across "types of crimes" and across "space": in fact we studied four different kinds of crime activities – murders, thefts, frauds and squeezes – in two years, 1999 and 2003, at the provincial level, so that we are able to capture important interregional disparities.

In Italy, in 1999 there were 4,171 crimes[1] per 100,000 inhabitants, in 2003 the value increased by 2.3%. This general measure of crime activities varies enormously in the "*Bel Paese*": in fact in north-western regions in 1999 this crime index was equal to 4,867, and decreased by 2.2% in 2003; in north-eastern regions this index was nearly equal to the Italian average value, 4,168, and increased by 4.8% in 2003; in central regions in 1999 the index was equal to 4,466, a value that increased by 4.3% in five years; in southern regions the value was the lowest, 3,383, but in 5 years it increased by 8.2%; and finally in islands, the index was 3,770 in 1999 and a 5.3% increased occurred in 5 years time. Concluding according to these statistics, Italy appears to be divided into North and South, with surprisingly no effects of the presence of crime activities related to organised crime. More plausible, as many general indexes, these statistics are missing lots of important features, in fact a more detailed analysis at "sectorial" level of crime reveals important "criminal" and "territorial" differences.

Maintaining the same territorial level of analysis (macro-regions), and limiting this preliminary analysis to crimes against person (murders) and crimes against patrimony (frauds) in 1999 and 2003, interesting *crimes geographies* emerge, increasing the suspicion of not homogeneous spatial distributions[2]. In 1999 in north-eastern regions, there were 2 murders per 100,000 inhabitants, and in the southern regions there were 7 murders (Italian average was equal to 4). In 2003, statistics and growth rates varied, but ranks remained similar: in north-eastern regions there was a nearly 9% increase (2.2 murders per 100,000 inhabitants), in the southern regions a dramatic decrease (18%), but the statistics still remained very high (6 murders against the Italian average value equal to 3.8).

Considering frauds, the picture varies dramatically: in 1999, in Italy frauds per 100,000 inhabitants were equal to 112, in north-western regions the value was almost double than the Italian

[1] This crime index (*indice di delittuosità*) is a gross measurement of crime activities and includes all kinds of crimes and provides a general intuition of crimes (Istat, 2006).

[2] Although statistics are different, thefts and squeezes show similar patterns to frauds and murders.

average (211), and in the southern regions the number of fraud was the half of the Italian average value (52). In 2003, in general there was an enormous increase of this kind of crime (191%) with 326 frauds per 100,000 inhabitants, and in the north-western regions the value remained very high (358 frauds per 100,000 inhabitants), while the islands registered the lowest value, 298, although the increase was dramatic: 451%! As we will emphasise later , this could be related to the kind of policy that the government followed during years, suspiciously augmenting the incentives to crime activities.

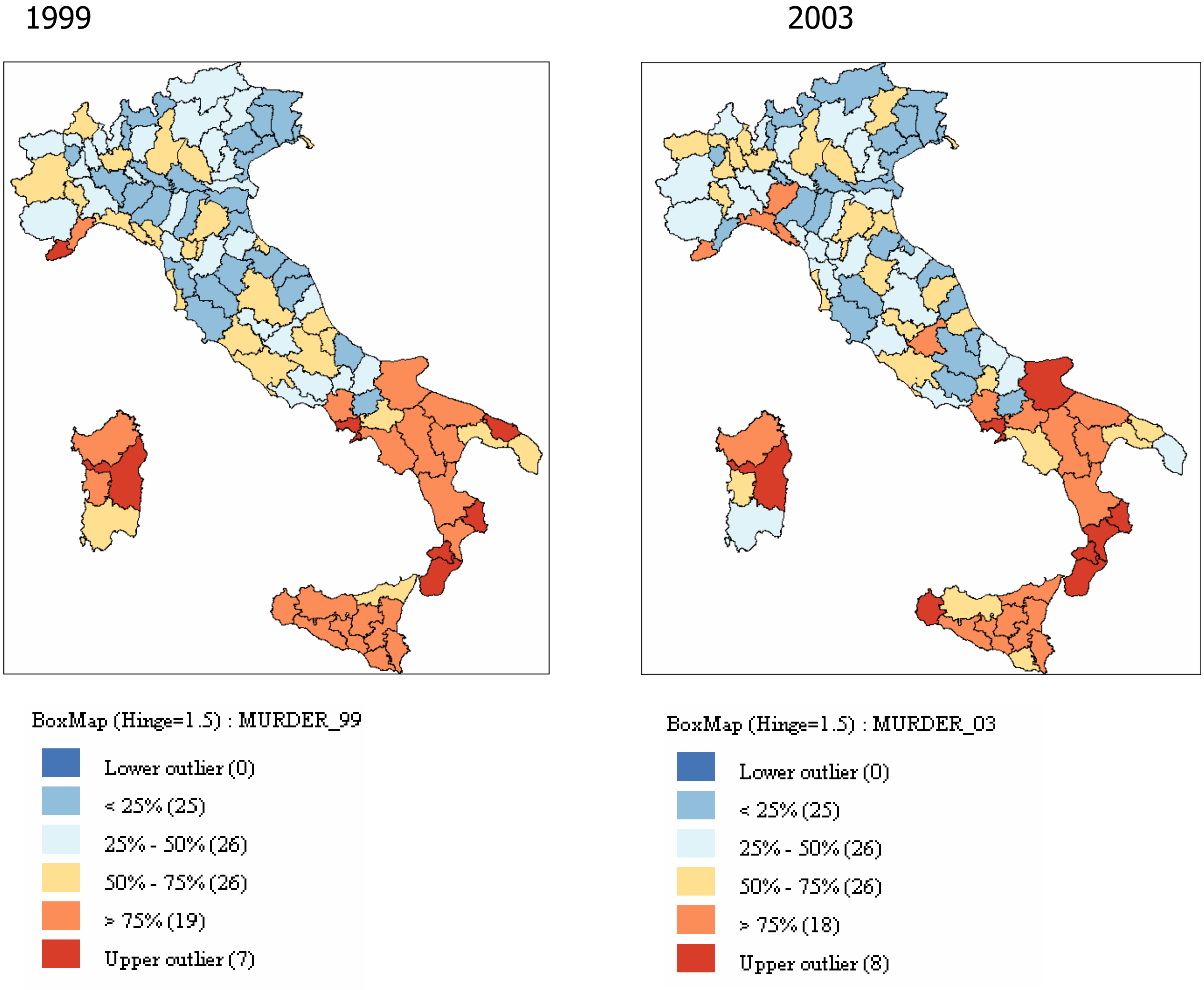
These aggregate statistics show that crime is a very complex phenomenon that needs to be investigated at appropriate disaggregate "territorial" and "sector" level of analysis (i.e. provinces and different types of offences) to catch important provincial disparities. Hence suitable tools of analysis (i.e. ESDA and spatial econometrics) should be used to control for possible spatial heterogeneity.

Hence in this analysis we decide to identify four kinds of crime activities: murders, thefts, frauds and squeezes, i.e. crimes against the person and against the property, and understand how organised crime, socio-economic conditions, presence of foreigners affect crime activities in the Italian provinces. The analysis is conducted for two years, 1999 and 2003, the first and the latest year available to conduct a comparable analysis and, most importantly, to detect, among other things, the effects of different policies on the criminals' incentives to offend.

The first step to identify possible patterns of spatial autocorrelation is to map the phenomena and conduct an exploratory spatial data analysis (ESDA). Here we present the box map of the distribution of crimes in the Italian provinces. This visual inspection is particularly useful because shows the location (quartile) of each province in the entire distribution of crimes, and allows to detect the presence of outliers. In this section firstly we show the ESDA for each type of crime, then we present the tests for spatial autocorrelation and finally the analysis for local spatial autocorrelation using LISA (Local indicator of spatial association).

As shown in Figure 1, the distribution of murders is very localised: in both years a clear cut line dividing southern provinces (e.g. Campania, Puglia, Basilicata, Calabria, Sicilia and Sardegna) from the rest of Italy exists. In all these provinces, organised crime is deepened in the territory, and a particularly bloody crime is spread. In fact as well known in the literature, this kind of crime is directly linked to the presence of these organisations. These maps seem to support the suspicion of a positive spatial autocorrelation that needs to be detected properly with appropriate indexes and tests (see Tables 1-3). The maps of murders in 1999 and 2003 are not very different, although in 2003, two provinces (Imperia and Savona) in the north-west part of Italy are outliers and show a certain degree of concentration of crimes. In 1999, there are 7 outliers: Nuoro in Sardegna, Napoli in Campania, Reggio Calabria and Vibo Valentia in Calabria, Brindisi in Puglia, and Imperia in Liguria. In 2003 there are 8 outliers, but exclusively in the South of Italy: half of them is located in Calabria (a region where the local organised crime, *'Ndrangheta* is very active).

Figure 1: Distribution of Murders in Italian provinces



Considering the second kind of crime, thefts (Figure 2), the distribution seems to be a more "Northern phenomenon", with the exception of Roma province. Inspecting the maps more in deep, it emerges that provinces affected by this crime are mostly "big cities" provinces (i.e. Bologna, Milano, Torino, Palermo) and some "tourist" provinces, like Rimini and Venezia. Besides this distribution reflects the Italian economic divide: in fact this crime seems to be associated with economic wealth (there exists a positive correlation between GDP per capita and thefts). We should also remember that this might be related to a different attitude in promptly denounce offences, especially for small kinds of thefts (i.e. pickpocket).

Finally in these maps thefts are clearly not a regional phenomenon, with the only exception of Liguria provinces, with a high number of this offence. This territorial heterogeneity justifies the importance to identify appropriate territorial levels of analysis, i.e. provinces, instead of regions. Differently from Figure 1, thefts do not show a clear spatial autocorrelation, hence we need proper statistics and tests to verify its presence (see Tables 1, 2 and 3).

Figure 2: Distribution of Thefts in Italian provinces

1999 2003

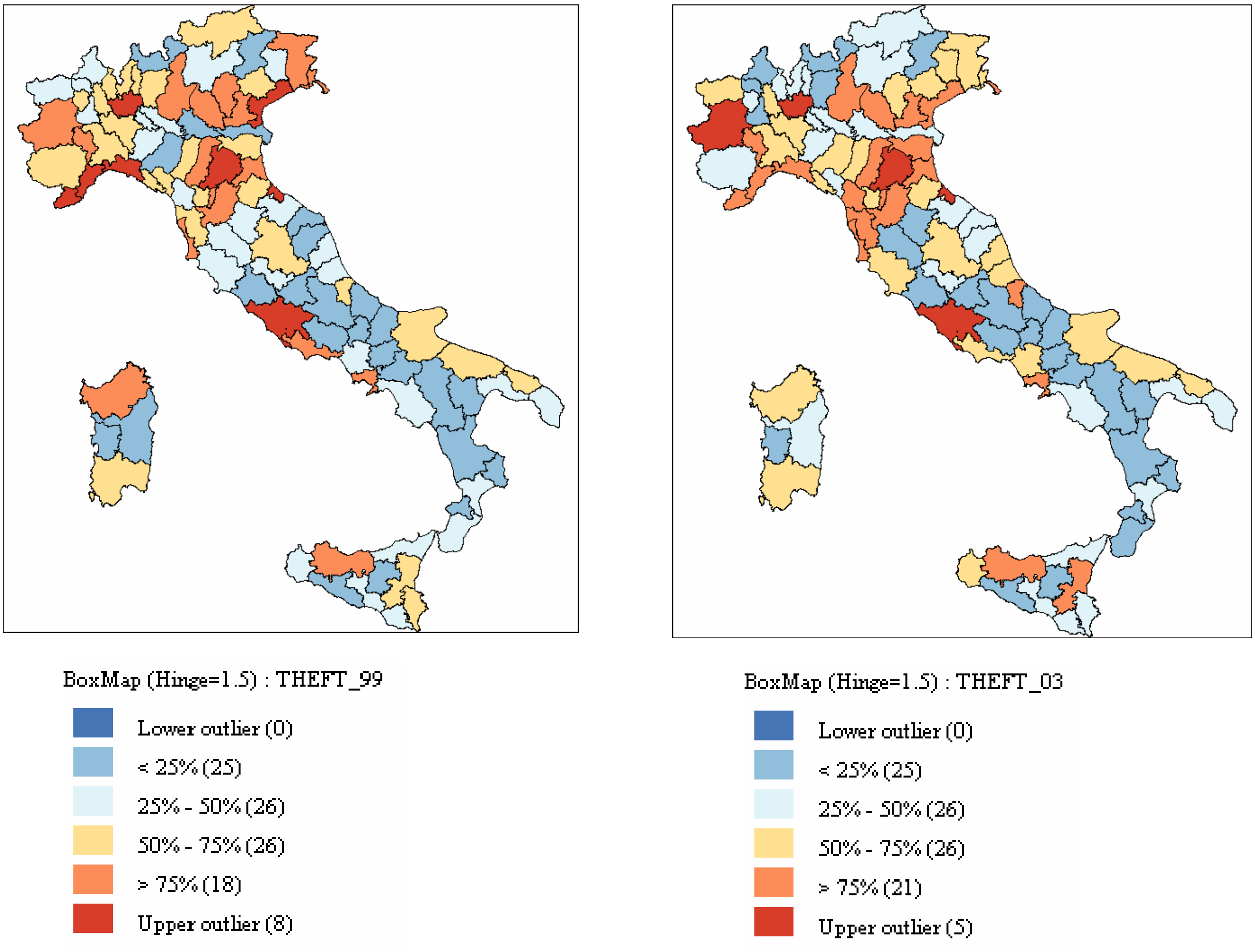
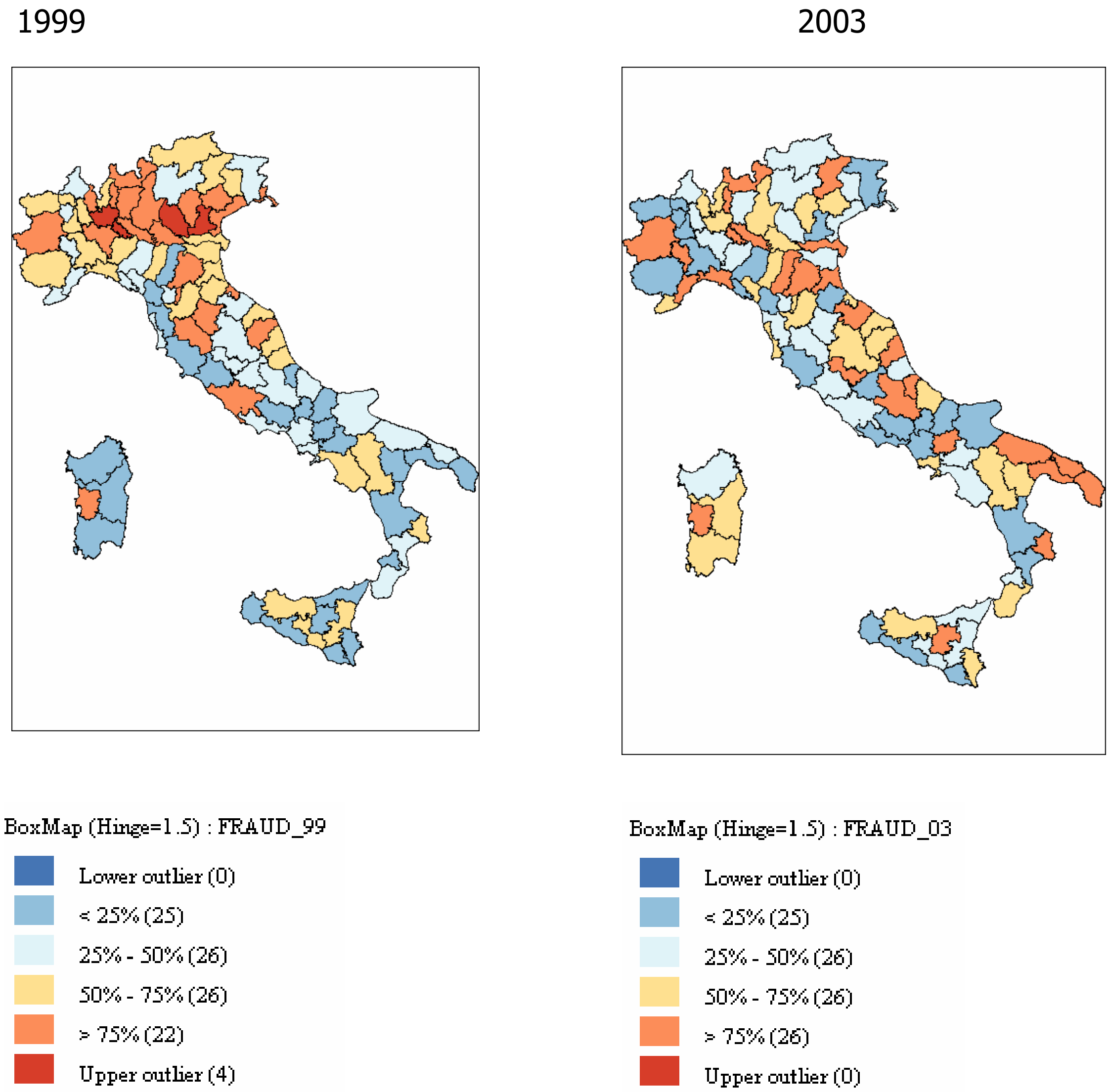


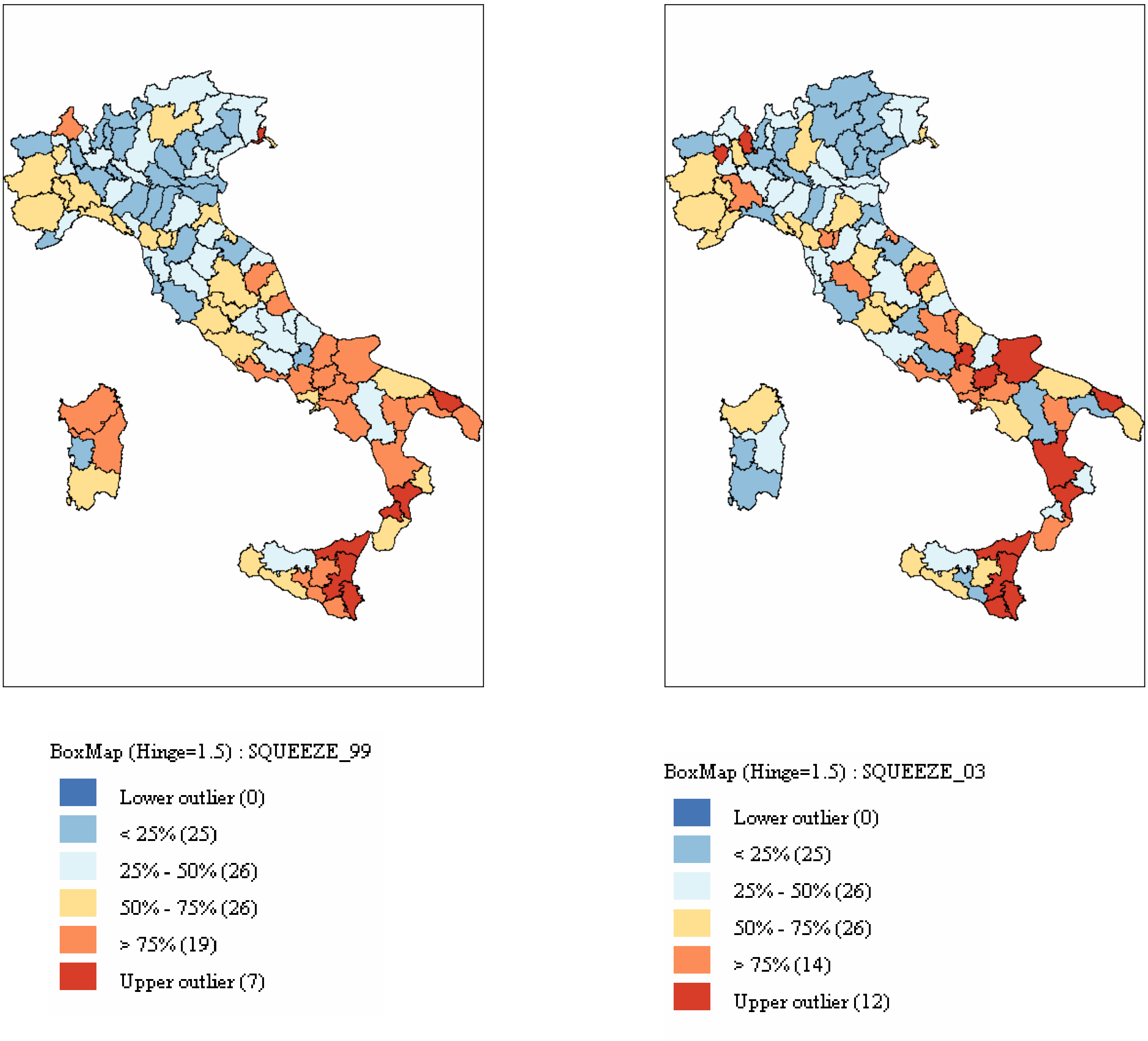
Figure 3: Distribution of Frauds in Italian provinces



Frauds spatial distributions constitute a very interesting phenomenon (Figure 3). Firstly in 1999, frauds are a North Italy phenomenon. There are only 4 outliers located in wealthiest parts of Italy (Milano and Lodi provinces in Lombardia; Padova and Verona in Veneto). In addition Lombardia seems to be the region most affected by this kind of crime, followed by the southern part of Veneto. Similarly to the previous kind of offence, in this case there exists a positive correlation with GDP per capita as well. Interestingly in 2003 the apparent positive spatial diffusion disappear abruptly, there are no outliers, the phenomenon of frauds spreads all over Italian provinces, but some spatial clusters appear in contiguous provinces (Savona, Genova, Asti, Torino; Sondrio. Lecco, Lodi Cremona; Bologna, Modena, Ravenna; Lecce, Brindisi, Taranto, Bari).

Figure 4: Distribution of Squeezes in Italian provinces

1999 2003



Finally, the box map of squeezes (Figure 4) reveal a distribution of crime activities very similar to the murders one. This result is probably related to the fact that squeezes need a particular "organised" criminal network structure, that is provided mostly by the organised crime, diffused in the southern regions. Interestingly outliers increase from 1999 to 2003. In 1999 most of the outliers are concentrated in eastern Sicilia and part of Calabria, with an outlier in Gorizia, a province nearby the Slovenian border. In 2003 the number of outliers increases, but apart from the eastern Sicilia and Calabria, other provinces are involved in the South (Benevento, Isernia, Foggia, Brindisi), and in the North (Varese and Biella) .

As shown by these maps, spatial autocorrelation seems to affect all kinds of crimes . Hence to detect the presence of spatial autocorrelation for different types of crime activities, we use Moran-I and Geary-C. The former captures the "global" spatial autocorrelation, i.e. if provinces with high crime rate are clustered nearby or not, and ranges between -1 and +1. The latter identifies the presence of spatial autocorrelation, and is used to describe differences in small neighbourhoods, if its value is less than 1 there is a positive spatial autocorrelation, if higher than 1 there is negative spatial autocorrelation.

Tables 1, 2 and 3 show the results relative to the Moran-I and Geary-C tests based on three different kinds of weights matrices: the contiguity matrix (W^C), the geographical proximity matrix (W^P) and the relational weights matrix (W^R). This latter matrix is build according to the migration flows among Italian provinces occurred in 1999 and 2001 (latest year available) to detect if the presence of social networks – being in contact with the original places of birth – do, or do not, play an influential role in the diffusion of Italian crime activities. The choice to include two different geographical weights matrices is mostly due to the intention to distinguish contiguity geographical effects (i.e. sharing a border) and distance based effects (e.g. in terms of travelling distance)

While the first two matrices capture spillovers effects, related to the geographical neighbourhood according to different perspectives, the relational matrix captures the possible social network effects (see also, Maggioni et al., 2007).

Respect to the contiguity matrix (Table 1), we immediately see that positive spatial autocorrelation exists for all kind of crimes, both using Moran-I and Geary-C statistics, but there are interesting differences for crimes and across time. First of all, spatial autocorrelation diminishes for all kinds of crimes: this could be due to the fact that in these two years, crime activities tend to spread all over Italian provinces, partially reducing the regional differences. In fact the number of crimes is not diminished in the period, but increased, while the coefficient of variation decreased. In 1999 murders and squeezes have the highest values of Moran-I (and lowest for Geary-C), while in 2003 these two types of crime behave differently: murders Moran-I is still high, although has a lower value, but squeezes Moran-I is very low, 35% of the previous value in 1999.

Theft is a type of crime not particularly affected by spatial autocorrelation: in both years is very low (0.1877 and 0.1174), suggesting that thefts are spread all over provinces, with no particular geographical peculiarity. Geary-C values confirm these results, with very high values nearly equal to 1.

Fraud is a very particular case, because its value changes abruptly from 1999, when there exists high and positive spatial autocorrelation (i.e. provinces with high number of frauds are clustered, see Figure 3), to 2003, when the spatial autocorrelation is absent (Moran-I and Geary-C are not significant). These interesting results might also be related be due to policies adopted in 2003: the so called *indultino*, a law that instituted the pardon law, was adopted in the August 2003 (legge 1 agosto 2003 n° 207), and this could have affected the incentives to crime activities all around Italian provinces.

Table 1: Moran-I and Geary-C based on contiguity matrix (W^c)

	Moran's I	Z Score	Geary's C	Z Score
Murders 1999	0.4842***	7.1832	0.5372***	-6.3997
Murders 2003	0.4446***	6.6067	0.5745***	-5.8839
Thefts 1999	0.1877**	2.8715	0.8173**	-2.5268
Thefts 2003	0.1174*	1.8494	0.8681*	-1.8233
Frauds 1999	0.3800***	5.6677	0.7649**	-3.2507
Frauds 2003	0.0379	0.6940	0.9470	-0.7336
Squeezes 1999	0.4682***	6.6495	0.5259***	-6.5560
Squeezes 2003	0.1656**	0.6940	0.8660*	-1.8526

Table 2 shows the Moran-I and Geary-C tests using the geographical proximity matrix, allowing to weight for geographical distances and detect for diminishing spillovers effects. Results confirm previous results obtained using the geographical contiguity matrix. All values decrease from 1999 to 2003 and in frauds the spatial autocorrelation disappears, and positive spatial autocorrelation is highest in murders and squeezes in 1999.

Table 2: Moran-I and Geary'-C based on geographical proximity matrix (W^P)

	Moran's I	Z Score	Geary's C	Z Score
Murders 1999	0.2178***	14.7937	0.7129***	-14.7489
Murders 2003	0.1738***	11.9342	0.7503***	-12.8282
Thefts 1999	0.0634***	4.7602	0.9292***	-3.6376
Thefts 2003	0.0575***	4.3769	0.9551**	-2.3047
Frauds 1999	0.1322***	9.2282	0.9289***	-3.6515
Frauds 2003	-0.0055	0.2776	0.9923	-0.3640
Squeezes 1999	0.2312***	15.6635	0.7085***	-14.9697
Squeezes 2003	0.0831***	6.0351	0.8753***	-6.4039

Table 3 reports the Moran-I and Geary-C tests using the relational matrix of migration and results show some ambiguities, especially comparing two tests. First of all, crimes are not always positively spatially autocorrelated: thefts are negatively spatially autocorrelated in 1999 and 2003 according to Geary-C and in 2003 this is confirmed also by a negative and significant Moran-I. Similarly frauds show some ambiguities: in 1999 they are affected by positive spatial autocorrelation, according to Moran-I and negative according to Geary-C, in 2003 spatial autocorrelation (both positive and negative) disappears completely.
In 1999 murders and squeezes are positively spatially correlated, but in 2003 the spatial correlation reduces and for squeezes Moran-I is not significant any longer (while Geary-C is significant).

Table 3: Moran-I and Geary-C based on relational proximity matrix (W^R)

	Moran's I	Z Score	Geary's C	Z Score
Murders 1999	0.0586**	2.31111	0.7940**	-2.6073
Murders 2003	0.0396*	1.6669	0.7880**	-2.5670
Thefts 1999	-0.0189	-0.3083	1.4716***	5.9700
Thefts 2003	-0.0691**	-2.0004	1.5953***	7.2075
Frauds 1999	0.1720***	6.1406	2.5437***	19.5433
Frauds 2003	0.0256	1.1958	0.8695	-1.5802
Squeezes 1999	0.0861**	3.2392	0.7536**	-3.1191
Squeezes 2003	0.0335	1.4627	0.8253**	-2.1158

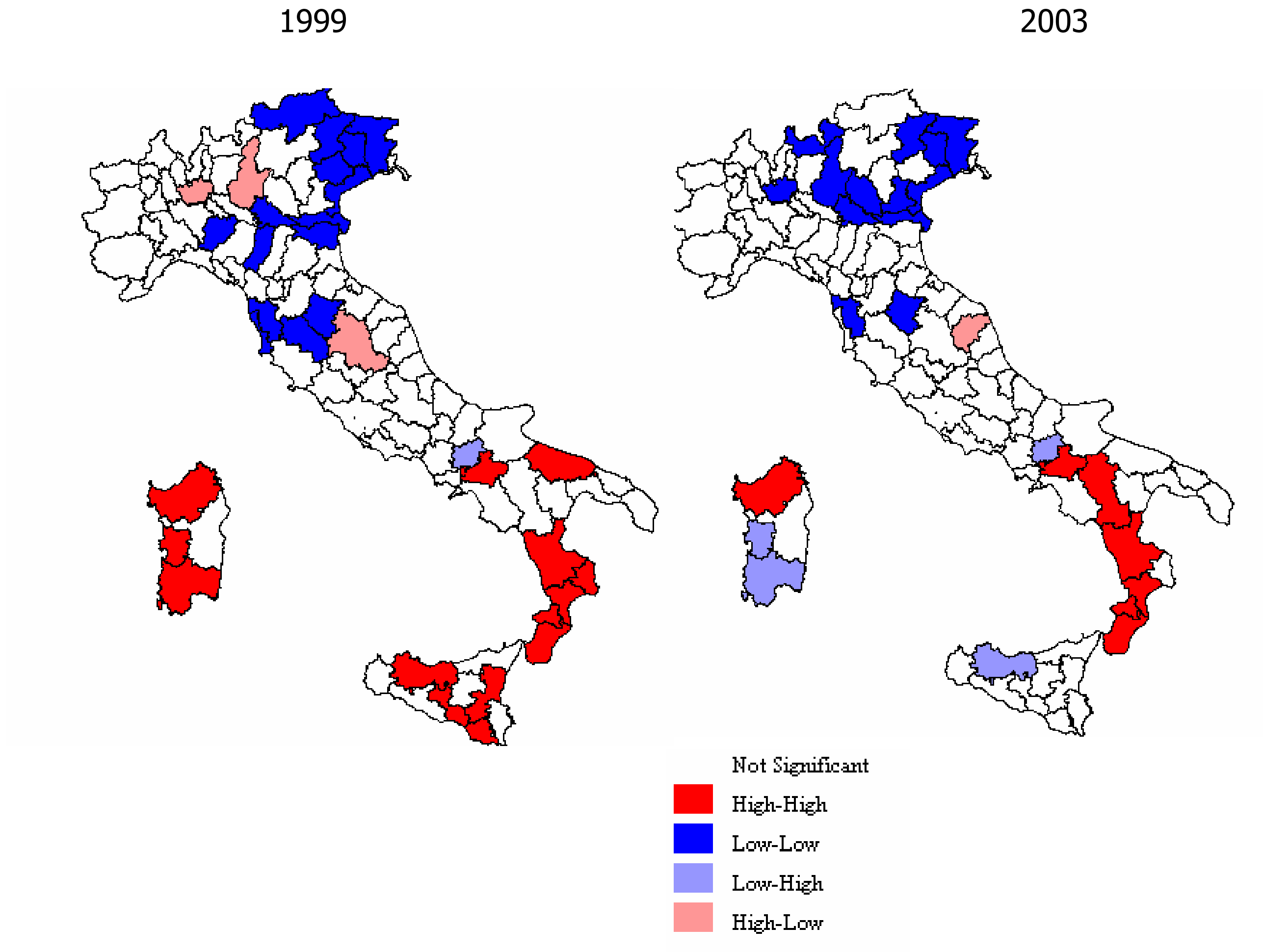
To conclude it seems that crime activities are related also to the "relational proximities", migration flows between provinces and not only to geography justifying an analysis that tries to include also this social networks effects.

Maps allow to detect the possible geographical patterns, Moran-I and Geary-C test for the presence of "global" spatial autocorrelation, but do not individuate local clusters. Hence to identify the local clusters we use the local indicators of spatial association (LISA)[3].

Figure 5 shows the LISA cluster maps for murders. It clearly emerges a clear cut distinction of two spatial clusters: those belonging to high-high cluster, with powerful organised crime, and those belonging to low-low cluster. Interesting spatial outliers, high-low and low-high combinations are represented by Perugia, Milano and Brescia, and Benevento. In 2003 an high-high spatial cluster emerges from Reggio Calabria to Benevento, while the low-low spatial cluster modifies its shape respect to 1999, including Brescia and Milano, previously belonging to spatial outliers.

[3] Because of similarities between the geographical distance and the proximities weights tests for spatial autocorrelation, here we show the LISA results for the contiguity matrix. It will be interesting to detect the LISA using the relational distance matrix, but Geoda, the software used to map the LISA, does not allow to use weights matrices unless they are dichotomised. Hence the relational matrix should be dichotomised according to a threshold value, above which there exists a "relational" proximity, otherwise not. The scope of this paper is firstly to analyse the effects of different weights matrices on crime activities, and we do not concentrate on the dichotomisation of the relational weights matrix. Hence for this reason we show only the LISA cluster maps for contiguity matrices.

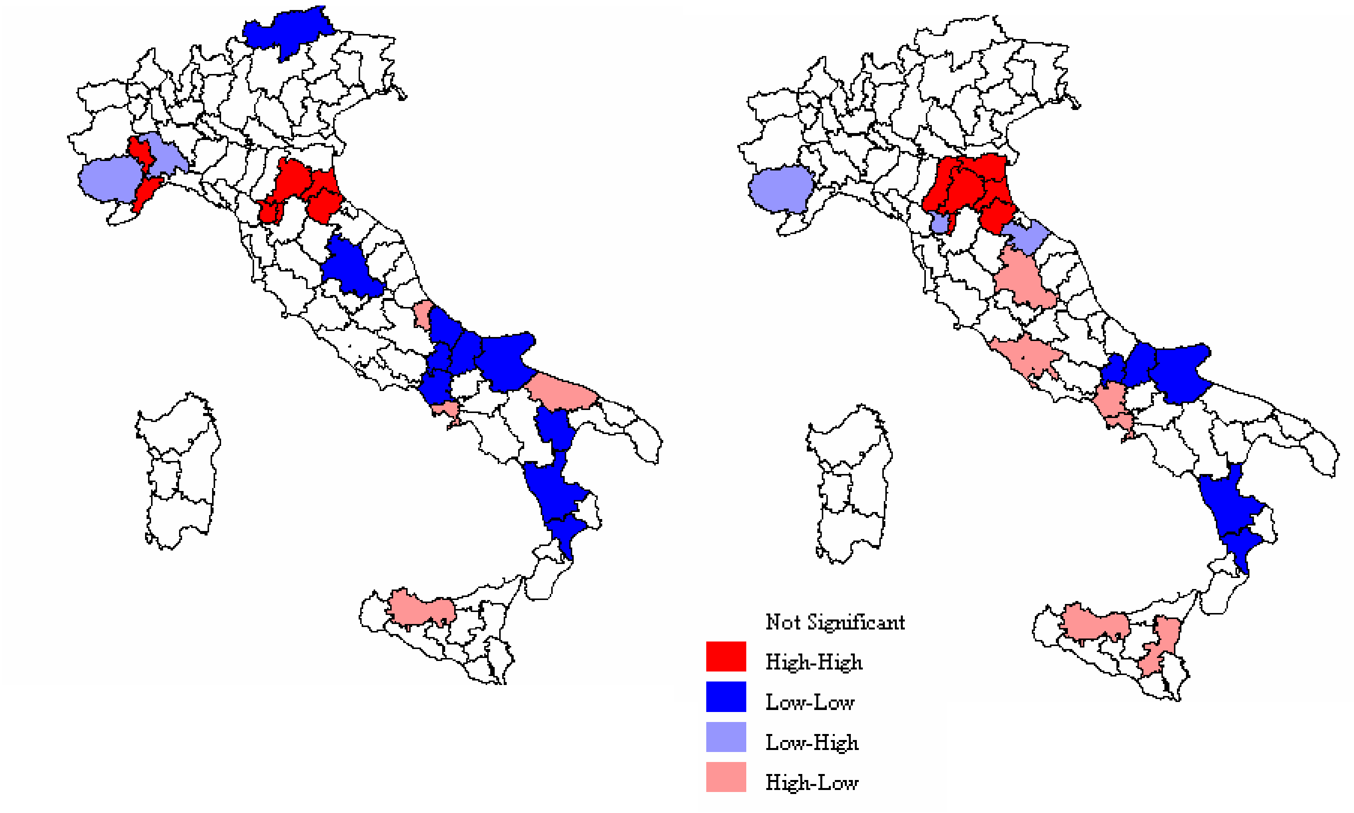
Figure 5: LISA cluster maps for Murders



Lisa cluster maps for thefts (Figure 6), especially for 2003, are ambiguous: in fact there are numerous spatial outliers belonging to low-high and high-low. Only exceptions to the presence of outliers are represented by the surrounding of Bologna provinces that belongs to high-high spatial cluster, Gargano region and Calabria belonging to low-low spatial clusters. These results could confirm the values of Moran-I that are very low (Table 1) and should affect the results of the analysis.

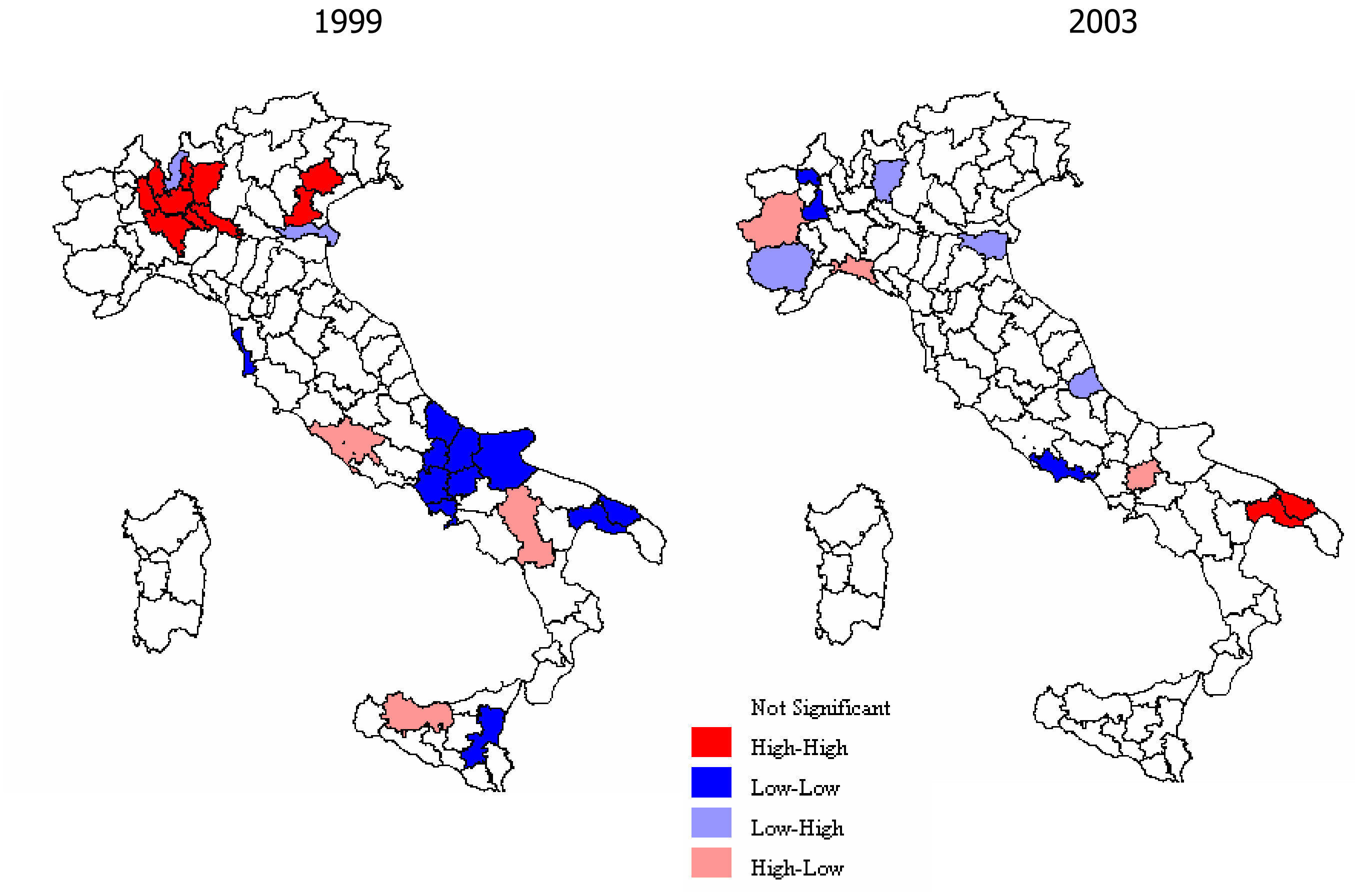
Figure 6: LISA cluster maps for Thefts

1999 2003



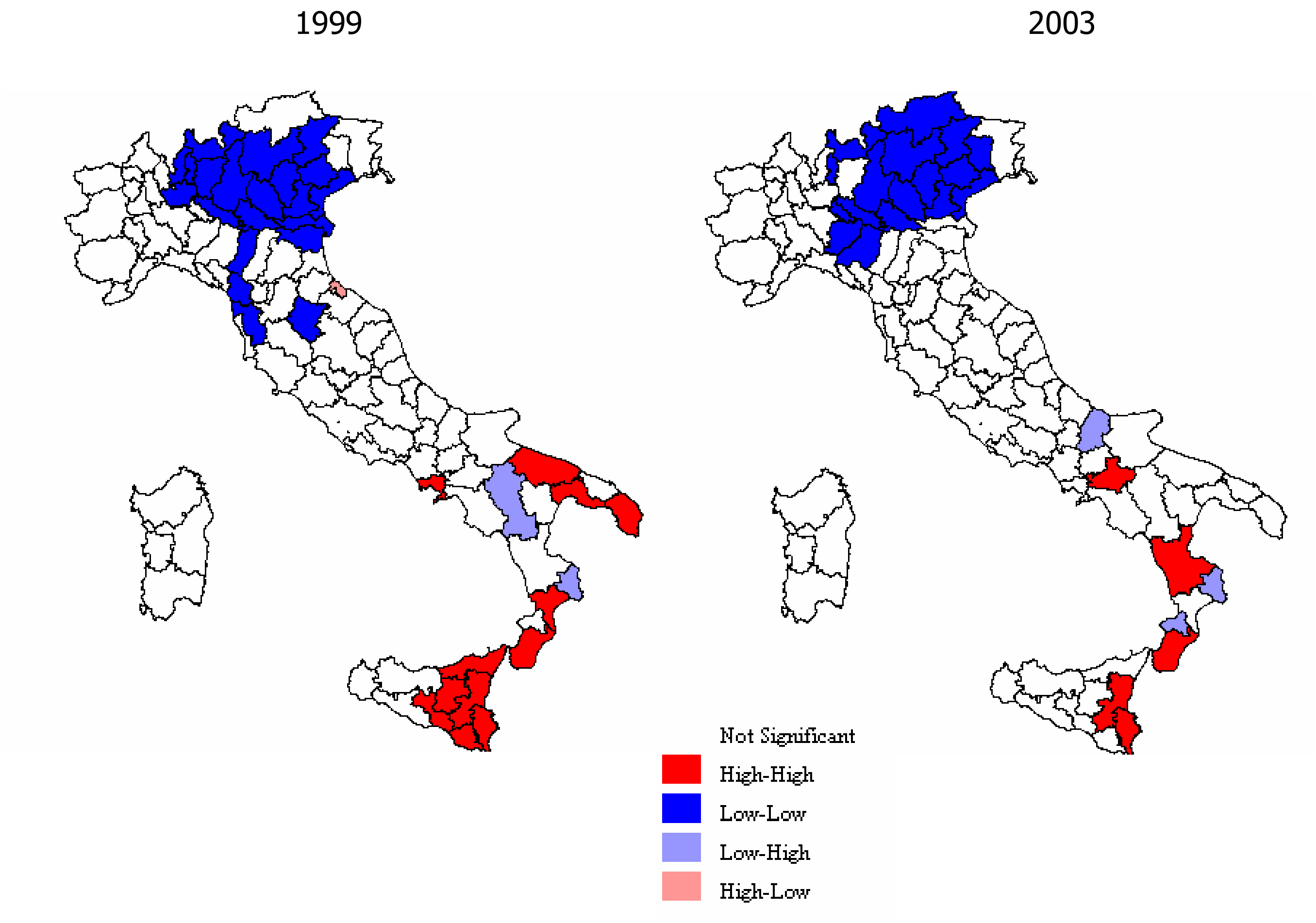
Lisa cluster maps for frauds (Figure 7) changes dramatically from 1999 to 2003. In 1999 there is an evident high-high spatial cluster involving most of the Lombardia region; and a low-low cluster that cut horizontally Italy from Tirreno to Adriatico Sea and in the southerner part of the "hill". Some outliers (both low-high and high-low) are spatially distributed along all the Italian peninsula. In 2003 there is a clear evidence of spotted spatial outliers suggesting for some heterogeneity problems in the empirical analysis with the southerner part of the "hill" that radically changed its nature: from low-low it becomes a high-high cluster.

Figure 7: LISA cluster maps for Frauds



Finally the last Lisa presented is relative to squeezes (Figure 8). Similarly to murders Lisa cluster maps, in this type of offence there is a clear distinction between two spatial clusters: a northern cluster (low-low) involving mostly the Triveneto region and a territorial appendix in Emilia Romagna, and a southern spatial cluster (high-high) involving half of Sicily and continues to the peninsula. The spatial outliers are 3 and spread around Italy. The picture slightly changes in 2003: the northern region belonging to the low-low spatial cluster becomes much denser, and the high-high spatial cluster become more sparse.

Figure 8: LISA cluster maps for Squeezes



Concluding this section relative to the ESDA, we can say that crime activities in Italian provinces are affected by spatial autocorrelation, that changes over time: the spatial autocorrelation is more relevant 1999 and decreases in 2003. The territorial level of analysis, provinces, allows us to detect for intra-regional disparities otherwise distorted. The spatial autocorrelation is substantially the same, considering the geographical proximity or contiguity, and shows similar patterns (especially for murders and squeezes) using the relational weights matrix, reflecting the social network relations between the place of birth and the place of actual residence of immigrants.

3. Theoretical background and empirical model

In the recent years, based on the traditional Becker-Ehrlich deterrence model, many empirical studies have been emerged in the economic literature as an extension of CEM. Specifically, *atheoretical* models have been motivated, on the one hand, by the increase of criminal activities, and on the other one, by the relevant socio-economic aspects that characterize the development of countries affecting the dynamics of crime; i.e., regional unemployment differences, income inequality, migration, high level of education, etc.

Marselli and Vannini (1997) – following the CEM interpretation of crime phenomenon – extended the factors that might influence the crime rates in Italy. They related crime rates to deterrence variables but also to socio-economic and demographic variables: real consumption per capita, unemployment rate, public works, the share of employed in the service sector, social security benefits, average monthly salary, the share of young male on the total population, the share of students that achieved the secondary and high school degree. They found that the probability to be convicted affects positively on crime activities, discouraging the criminal activities and its effect is more effective than the severity of punishment and detective efficiency of judicial system. Further, among the socio-economic variables, unemployment, employment in service sector and number of public workers have a significant effect in explaining crime rates, while the level of education of population and the share of young males do not have any effect on crime rates relating to the different kinds of crimes.

Ehtorf and Spengler (2000), using a panel of the German Laender, tested the deterrence hypothesis by CEM, but also the effect of some socio-demographic and economic variables to explain the crime differences. They found that the deterrence hypothesis is verified, though its power is weaker with respect to crime against the person than crime against property. With respect to demographic variables, they found the young unemployment and the share of foreigners have a significant influence on the explanation of crime. Specifically, relating to the latter, it is positively associated with the crime against the property.

Using Johansen cointegration techniques, Luiz (2001) analyses the association between per capita crime levels and income per capita as proxy of economic opportunities, police officers per capita, conviction rates, and political instability. The analysis concerned different crime series for South Africa for the period between 1960 and 1993.

In addition to the aforementioned literature, it is worthy to notice there is a wide number of empirical studies focussed on the relation between crime disparities and the features of labour market and skills of working age population (see, e.g., Lochner 2004; Burdett et al. 2004; Edmark 2005).

Following the previous literature we propose a cross-sectional analysis of provincial crime rates in Italy for years 1999 and 2003. Considering the geographical structure of Italian crime activities, we develop a model including deterrence and socio-economic variables, as well as spatial dependence effects[4] to investigate the determinants of crime in Italy. Following Cherry and List (2002), to avoid possible bias estimation linked to aggregation of crime activities, we consider four different types of crime: murders, thefts, frauds and squeezes.

To explore both the significance of spatial clusters of crime rates and the determinants of crime activities, our starting point is a cross-sectional regression model on provincial crime rates without spatial effects. In particular, the following general theoretical model is used as a starting point:

$$\begin{aligned} C_j^t = \beta_o + \beta_1 Severity^t + \beta_2 Probability^t + \beta_3 Unknown^t + \beta_4 Sser^t + \beta_5 Sind^t + \beta_6 Umale_{25-29}^t + \\ + \beta_7 GDP^t + \beta_8 RGDP^t + \beta_9 Young^t + \beta_{10} Old^t + \beta_{11} Foreigners^t + \varepsilon \end{aligned} \tag{1}$$

where C_j indicates the different kinds of crime; *Severity* is the average time spent in prison as a proxy for deterrence variable; *Probability* is the share of offenders convicted over the total provincial reported offenders[5] as a proxy for deterrence variable; *Unknown* is the share of crimes committed by unknown persons over the total recorded crimes in each kind of crime as a proxy for deterrence variable. The expected sign of the latter deterrence variable is positive; while severity and probability have a negative expected sign, a longer average convict time and a high probability to be convicted should not incentive potential offenders to do illegal activities.

[4] At this first attempt we consider only the spatial dependence effect and not also the heterogeneity spatial effect.

[5] The data on severity and probability for each kind of crime are not available at provincial level, so we use the average time spent in prison and the offenders convicted relate to the all crime activities.

The other explanatory variables are proxies for socio-economics and demographic aspects. *Sind* is equal to the share of firms in the industrial sector over total provincial firms. *Sser* is equal to the share of firms in the service sector over total provincial firms. *Sind* and *Sser* are proxies for the provincial economic structure, though it is not always clear which sign these control variables should have. Clearly, intuitively, in provinces characterizing by high share of industrial and services firms it should be easier to find a job, so the inclination to commit an offence should be lower. $Umale_{25\text{-}29}$ is the unemployment rate of the cohort males from 25 to 29 years old, it is a proxy for opportunity cost of legal and illegal activities; a positive sign of its coefficient should indicate that people excluded from labour market as they do not have an income, they tend to commit a crime. *GDP* is the gross domestic product per capita as proxy for legal and illegal income opportunity; the excepted sign of the coefficient is negative. *RGDP* is the relative gross domestic product as share of the provincial GDP per capita over national GDP per capita as proxy for legal (or illegal) income opportunity, the excepted sign of the coefficient is negative. *Young* is the percentage of population between 19 to 25 years old over the total population; *Old* is the share of people over 65 years on the total provincial population. The expected sign of these coefficients are positive and negative, respectively; it is plausible to suppose that young people are more favourable to commit a crime for different social reasons: to follow peer group; the absence of reputation feeling, etc (e.g., see Eide 1994; Glaeser et al. 1996; Akerlof 1997). Finally, *Foreigner* is the share of resident foreigners over the total provincial population. In a country as Italy, characterized in the last years by high immigration rate, mostly from non-European countries it is correct to include such a variable for two main reasons: firstly because many empirical analyses demonstrated the influence of this variable on crime activities and secondly because it is a matter of facts that recent immigrants show an attitude to commit crimes similar to residents[6].

If spatial dependence effects are included in the model (1), this leads us to the concept of the spatial interaction between economic phenomena. This introduces to the concept of spatial autocorrelation, which is linked to the territorial shape of the observed phenomena and to the connections between observations. Measures of spatial autocorrelation take into account the dependence between observations by a spatial weights matrix **W**. For a set of *N* observations the spatial matrix **W** is an *N* x *N* matrix with diagonal elements equal to 0; the other elements w_{ij} represent the intensity of the effect of territorial area *i* on territorial area *j* (see Anselin and Bera 1998). The matrix defines the structure and the intensity of spatial effects, and it may be a contiguity matrix, or a proximity matrix based on a distance function, or a relational weights matrix. In the literature, there are very few formal guidelines and suggestions on the choice of the most adequate spatial weights (for details, see Anselin 1988, 2002; Anselin and Bera 1998; Leenders 2002; Dietz 2002). Here, we use three different kinds of matrices: rook contiguity and distance matrix, and a relation matrix based on the inter-provincial migration flows. All the matrices are row-standardized. The rook contiguity matrix is a binary spatial weight such that $w_{ij}^s = w_{ij} / \sum w_{ij}$ if the provinces *i* and *j* are contiguous (i.e., share a border), and $w_{ij}^s = 0$ otherwise. With respect to distance matrix, given the Euclidean distance between two provinces d_{ij}, each element of the matrix is equal to $w_{ij}^s = \dfrac{1/d_{ij}}{\sum 1/d_{ij}}$ for $i \neq j$, and $w_{ij}^s = 0$ otherwise. Finally, relating to relation weight matrix, given the migration inter-provincial flows, f_{ij}, the spatial weight is equal to $w_{ij}^s = f_{ij} / \sum f_{ij}$ for $i \neq j$, and $w_{ij}^s = 0$ otherwise.

As argued above, the three matrices let us to catch different spatial effects as contiguity, geographical distance and relational effects.

6 Recently the Home Office Minister, Giuliano Amato, presented a report on crime activities in Italy showing that offences attributed to foreigners increased rapidly during last 5 years.

The most general spatial model includes spatial dependence effects relating to dependent variable, independent variables and error terms. To identify the best specification of the model we follow the hybrid specification strategy (see Anselin et al. 1996; Florax et al. 2003) without ignoring the theoretical arguments on the basis of which model (1) was performed.

Considering the following spatial dependence models:

$$C = \rho WC + \beta X + \varepsilon; \tag{2}$$

$$C = \beta X + \varepsilon; \quad \text{with} \quad \varepsilon = \lambda W\varepsilon + \xi \tag{3}$$

where WC is a spatially lagged dependent variable, ρ is the spatial autoregressive coefficient and measures the spillover effect connected to the dependent variable; $W\varepsilon$ is a spatially lagged error term, λ is the spatial autoregressive coefficient and ξ is the error term[7].

According to the hybrid specification strategy, we test the statistical significance of λ and ρ, departing from a model without spatial effects using a separate LM test. We test whether λ and ρ are equal to 0; if neither are equal to 0, we could choose between a spatial error or a spatial lag model, on the basis of the largest robust LM statistics (see Anselin et al. 1996). If only LM_λ (or LM_ρ) is significant, a spatial error (or spatial lag) model could be estimated.

To test our model, data from different sources have been used: Murder, Theft, Fraud and Squeeze from Italian Statistic Agency, Judicial Statistics (ISTAT 1999a, 2003a), Severity, Probability and Unknown from Italian Statistic Agency, Judicial Statistics (ISTAT 2000, 2003a); Old and Young population from Italian Statistic Agency, Demographic Statistics (ISTAT 1999b, 2003b); Foreigners from Italian Statistic Agency, Territorial Indicators (ISTAT 1998, 2002); young male unemployment rate from Italian Statistic Agency, Italian Survey on Labour Force (ISTAT 1999c, 2003c); and value added per capita, Italian Statistic Agency, Regional Accounting (1999d and 2003d).

The empirical findings, discussed in the next section, were obtained in the light of the aforementioned specification strategy, theoretical and empirical arguments.

4. Empirical results

The previous section has identified – on the basis of theory, the empirical insights and the availability of data – some relevant variables that explain the differences of crime in Italian provinces. Clearly, it is not expected that all variables included in model (1) would be required in an adequate statistical model. In fact, in our case the estimation of model (1) has produced relevant statistical problems, like heteroskedasticity and multicollinarity. Hence, these problems had to be solved by testing a double log model, and following a model selection strategy. Specifically, we did not include in the statistical model the variables *Probability*, *Sser*, *GDP*, *RGDP*, *Young*, *Old* e *Foreigners*, because they either caused a severe multicollinearity problems or were not significant. With respect the last variable, we include in the model the foreigners at *t-1*; it is plausible to

[7] It is worth asking whether a more general model than the models 2 and 3 would be preferable; i.e. a model with spatially lag dependent and explanatory variables. In this case, an LR test on a common factor hypothesis should be done. Specifically, the autocorrelated error model is equivalent to a special form of spatial lag model by the following transformation of dependent and independent variables: (Y - λWY) e (X - λWX); so the spatial lag model can be written as Y = λWY + Xβ - λWXβ + ε. This is a subset, known as the common factor hypothesis model, of the more general model Y = λWY + Xβ + WXδ + ε. The LR test of the common factor hypothesis tests the hypothesis δ = λβ: if the null hypothesis is rejected a more general model with lagged independent variables must be estimated.

hypothesize that crime at the time *t* depends on the share of foreigners at *t-1*, in fact 1 year is a mean time during that the foreigners look for a job or an alternative illegal activity. That let us to eliminate correlation problem between foreigners and *Sind* (or $Umale_{25\text{-}29}$).

The estimable statistical model used for our final estimation is thus the following:

$$C^t = \beta_o + \beta_1 Severity^t + \beta_2 Unknown^t + \beta_3 Sind^t + \beta_4 Umale^t_{25-29} + \beta_5 Foreigners^t + \varepsilon \qquad (4)$$

Tables 4-7 show the OLS and ML estimates for the different kinds of crimes for 1999 and 2003. The results show that in 1999 crimes activities in Italian provinces are affected by spatial effect, on the contrary in 2003 this effect disappears with the only exception of thefts. Moreover, the expected signs are generally confirmed for deterrence and socio-economic variables. Further, the use of different weights matrices lead to catch the different strength of the spatial effect of crime activities.

With respect the murder activity for the 1999 and 2003, all the variables, with the exception of *Unknown* and *Sind* are significant. In 1999, in contrast to 2003, the LM tests on omitted spatially lagged dependent variables and spatially lagged error term by contiguity matrix appear to give a significant value equal to 9.48 and 3.09, respectively. Further, in 1999, relating to the proximity matrix the LM test on omitted spatially lagged dependent is also significant and equal to 6.37. While, using the relational weight matrix, no spatial dependence exists; it was expected given the low value of Moran-I statistics.

The significant value of the robust LM_ρ tests, for both contiguity and proximity matrix indicates that $\rho \neq 0$; so, a spatial lag model has to be estimated. The estimations of our spatial lag regression model are shown in columns 2-3 of Table 4. The estimations related to both matrices show a significant and consistent spatial effect explaining the differences of provincial murder activity. The coefficient of the variable WC_{murder} is rather large; $\boldsymbol{\rho}$ is equal to 0.33 and 0.67 relating to contiguity and proximity matrix, respectively (Table 4). The positive value of $\boldsymbol{\rho}$ implies that murders in province *i* depends directly on the murders in other neighbouring provinces; in other words provinces with high, or low, percentage of murders are clustered together. This result reflects the geographic distribution of this crime characterized by low and high rates of murders in the central-northern and southern provinces, respectively. In particular, the high rate of murders in the southern provinces could be connected to the phenomenon of organized crime that features many southern areas (e.g. *Camorra*, *Sacra Corona Unita*, *'Ndrangheta*, *Mafia*, etc.). Moreover, the murder differences are explained by the *severity*, while the unknown is not significant in both spatial lag models. The positive sign of the coefficient of *Severity* implies that percentage of murders depends strongly on the severity of punishment, i.e. an increase of one year of the punishment gives a more than proportional increase of murders. This result, in contrast to the many empirical findings from the literature, could indicate a vicious circle; i.e. a person that spend many years in prison, once he/she served a long sentence it will be difficult to enter again in the 'society', to look for an job and to find an job because he/she does not have adequate social network knowledge and skills to change his/her living; so he/she will enter easier into local crime groups ('*peer group*') because these are 'settings' he/she know very well (see Homant 1984; Orsagh et al. 1988). All that may lead released people from the prison, with a low attitude to change his/her living, to start again to commit a crime as consequence of his/her entrance in the crime groups. On the other hand, the positive relation between murder and severity might be connected to a high attitude to violent behaviour as the consequence to spend a long period in prison; i.e. a person that spend many years in prison, that is a environment characterized by violent and repressive behaviours, could improve his/her violent attitude becoming inclinable to do crime activities (Eronen et al. 1996). In 1999, moreover, the murder activity is positively linked to the young male unemployment; in other words, in the provinces where the murder rate is high, the young male unemployment is high. This result reflects the geographical distribution of murder and unemployment in Italian provinces: high (or low) rate of murders and young unemployment together characterizing the southern (or northern) provinces. Finally, the *Foreigners* variable show a lower strength to explain the provincial murder differences with respect to the others variables. Even if the coefficient is weakly significant its positive sign indicates that a marginal increasing of resident foreigners gives a less proportional increase of murders (Table 4).

In contrast to 1999, in 2003, although the Moran-I statistics is positive and significant, the OLS diagnostic tests do not show any spatial dependence for the murder activity. Similarly to 1999 (OLS estimation column 1), in 2003 only the coefficient of severity and young male unemployment are significant. The latter has a value almost equal to that of 1999, while the strength of severity is smaller; i.e., in 2003, the severity effect declines, this might be connected to a higher attitude, with respect to 1999, of released persons to improve or change their life.

Relating to thefts , for both years all the explanatory variables are significant with the exception of *Unknown* and $Umale_{25\text{-}29}$ variables that in 2003 are not significant (Table 5). In 1999, the diagnostic for spatial dependence indicates a spatial autocorrelation exists and it pertains to the dependent variable, if we use a proximity matrix, or the error term, if we use a contiguity and relational matrix. The estimations obtained by the three different spatial models do not show relevant differences in terms of sign and significance of coefficients. So, we focussed on model in column 4 Table 5 because it presents the lowest AIC value. In 1999, the thefts are affected by a high contagious effect equal to 0.816 and all explanatory variables are significant with the exception of $Umale_{25\text{-}29}$. It is worthy to notice both deterrence variables have a positive effect on theft, though the efficiency of police authority in deterring crime (*Unknown*) is more effective than the severity of punishment. As said above, relating to the positive effect of severity on theft, this reflects *vicious circle* where a higher punishment gives to a higher attitude to commit crimes. Also in 2003, the theft crime is characterized by a contagious effect that is catch by a contiguity weights matrix, but in contrast to 1999 the *Unknown* is not significant. It is worthy to notice that in contrast to the other types of crime, the *Sind* is significant in both years and its negative effect on theft is higher in 1999 than 2003.

Finally, the estimations related to fraud and squeeze activities show a very different performance in the two years analysed. In particular, in 1999 relating to fraud activity, all variables with the exception of *Sind* are significant. The significant values of the LM_ρ test and robust LM_ρ one relating to three spatial weights matrices indicate that a spatial lag model has to be estimated. The ML estimates relate to all spatial weights matrices do not show relevant differences in term of significance and sign of the coefficients. Therefore, we focus on the model in column 2 obtained by using a contiguity weights matrix because it is the model that satisfied the ordering of Wald (W), Likelihood Ratio (LR) and Lagrange Multiplier statistics in terms of their magnitude; i.e. W>LR>LM; also AIC value has been also considered to chose the best model. In particular, in contrast to murders, only the *Unknown* variable is significant and its coefficient has the positive expected sign. Therefore, the fraud crime is not sensible to the deterrence variable as *severity*, while the *unknown* deterrence variable has almost proportional effect on frauds discouraging its activity. Among the socio-economic variables, $Umale_{25\text{-}29}$ and *Foreigners* present significant coefficients. In particular, the relevant value of the *Foreigners* coefficient should lead policy makers to pay attention on the effects linked to the presence of migrates on crime activities as frauds. The significant and positive sign of the variable confirm the insights provided by the descriptive statistics from ISTAT (1999a) that underlined that foreigners have a similar attitude to resident population to commit a crime. Further, the fraud crime is characterized by a high contagious effect equal to 0.58. This means some policies focussing on the reduction of fraud in the province *i* could give a positive effect on the neighbouring provinces.

In 2003, in contrast to 1999, the frauds OLS model does not describe well the phenomenon; in fact all variable, with the exception of the constant, are not significant and the R^2 test on goodness fit is very low (0.03). The low value of the R^2 test indicates that some other variables should be included in the model to describe the frauds difference in 2003; the bad fit of the model lead to hypothesize that some relevant or 'shock' effect have acted on the featuring of frauds. This 'shock' effect might be the new regulation of '*indultino*' applied in 2003; the law provided a discount of the punishment for some kinds of crime as fraud. It is plausible to think that punishment discount encouraged people to commit crime activities. This immeasurable effect – that gets to reflecting in the high increasing of fraud offences in 2003 respect to 1999 – could explain a part of the bad fit of the model. Moreover, other causes could explain the fraud offences and improve the fit of the model. Maybe, other relevant variables should be included in the model in order to take into account the socio-economic differences between the 1999 and 2003, like the percentage of employees in the service sector as a proxy of fraud offences connected to white-collar workers; in fact, these are strongly linked to the criminal organization. Because of, severe multicollinearity problem, the number of employees in service sector has not been included in our analysis. In order to catch the effect of this variable further improving could be done by using instrumental variable methods.

Similar considerations could be done with respect to squeezes for the 2003, in fact only $Umale_{25\text{-}29}$ presents a significant coefficient (Table 7). While in 1999, the squeeze crime is explained by the $Umale_{25\text{-}29}$ and *Unknown*, both variables show the expected sign. Further, the diagnostic on spatial dependence indicates that a spatial lag model for both contiguity and proximity matrix has to be estimated. The best spatial lag model was the model in column 3 Table 7 related to the proximity matrix and with the lowest AIC. The spatial lag model shows that squeezes are affected by a relevant and significant neighbouring effect equal to 0.74. Further, between the deterrence variables only the *Unknown* has a positive effect on squeezes; while among the socio-economic variables only the $Umale_{25\text{-}29}$ has a positive and less than proportional effect on squeeze crime.

Briefly, the analysis shows a neighbouring effect characterize the crime activities in 1999, while this effect disappears in 2003. For all crime activities, with the exception of murder , the efficiency of police authority in deterring crime is more effective than the *severity* of punishment. Among the socio-economic variables, *Sind* does not affect on the explanation of crime differences with the exception of theft crime; while $Umale_{25\text{-}29}$ and *Foreigners* have a significant effect for each crime and for both years analysed.

Finally, the bad fit of the empirical model for 2003 leads us to investigate in-depth on appropriate variables able to explain the differences of crime in Italy. As the main deterrence and socio-economic variables presented in literature were included in our empirical model the results obtained should recall the attention on the need to define more sophisticated model including some raised socio-economic changes that could act on crime as stress, time dedicated to leisure activity, less health style of life, socio-economic well-being inequality, high female unemployment rates; high level of education of female population; etc.

Table 4: OLS and ML for Murders, years 1999 and 2003

Independent Variables	Murder^{99}			Murder^{99} W^C	Murder^{99} W^P	Murder^{03}		
	OLS (1)			Spatial model (2)	Spatial model (3)	OLS (4)		
Severity	1.243**			1.087**	1.037**	0.864**		
	(0.002)			(0.002)	(0.004)	(0.002)		
Unknown	0.162			0.111	0.134	0.135		
	(0.140)			(0.267)	(0.190)	(0.105)		
Sind	-0.283			-0.332	-0.290	-0.023		
	(0.333)			(0.213)	(0.288)	(0.923)		
Umale$_{25-29}$	0.443***			0.323***	0.324***	0.385***		
	(0.000)			(0.000)	(0.000)	(0.000)		
Foreigners^{t-1}	0.099			0.174*	0.155*	0.007		
	(0.314)			(0.052)	(0.091)	(0.924)		
Constant	0.853			0.747	0.348	0.208		
	(0.340)			(0.359)	(0.679)	(0.795)		
ρ	-			0.334**	0.670***	-		
				(0.001)	(0.000)			
AIC	203.990			196.869	200.128	164.464		
Log Likelihood	-95.995			-91.435	-93.064	-76.232		
	W^C	W^P	W^R			W^C	W^P	W^R
MORAN (Error)	2.400**	2.588**	2.031**			1.032	1.081	-0.013
	(0.016)	(0.010)	(0.042)			(0.302)	(0.280)	(0.990)
LM Lag	9.484**	6.374**	1.117			2.033	0.769	0.000
	(0.002)	(0.012)	(0.291)			(0.154)	(0.383)	(0.997)
Robust Lag	7.496**	7.525**	0.010			3.803*	2.390	1.672
	(0.006)	(0.006)	(0.922)			(0.051)	(0.122)	(0.196)
LM Error	3.904**	0.856	1.475			0.368	0.010	0.197
	(0.048)	(0.355)	(0.225)			(0.544)	(0.920)	(0.658)
Robust Error	1.917	2.007	0.367			2.139	1.640	1.868
	(0.166)	(0.157)	(0.545)			(0.144)	(0.200)	(0.172)
LR Test Lag				9.121**	5.862**			
				(0.003)	(0.015)			
LM Test Error				2.283	0.018			
				(0.131)	(0.892)			

Table 5: OLS and ML for Thefts, years 1999 and 2003

Independent Variables	$Theft^{99}$	$Theft^{99}$ W^C	$Theft^{99}$ W^P	$Theft^{99}$ W^R	$Theft^{03}$	$Theft^{03}$ W^C
	OLS (1)	Spatial model (2)	Spatial model (3)	Spatial model (4)	OLS (5)	Spatial model (6)
Severity	0.396**	0.387**	0.444**	0.316*	0.509***	0.590**
	(0.043)	(0.028)	(0.15)	(0.074)	(0.000)	(0.001)
Unknown	8.533***	8.644***	8.273***	8.238***	-0.327	0.049
	(0.000)	(0.000)	(0.000)	(0.000)	(0.533)	(0.923)
Sind	-0.400**	-0.410**	-0.394**	-0.447**	-0.647***	-
	(0.005)	(0.003)	(0.003)	(0.002)	(0.000)	0.734***
						(0.000)
$Umale_{25-29}$	-0.092**	-0.033	-0.038	-0.031	0.059	0.111
	(0.050)	(0.514)	(0.403)	(0.512)	(0.433)	(0.149)
$Foreigners^{t-1}$	0.257***	0.287***	0.246***	0.321***	0.255***	0.268***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	8.820***	8.644***	3.641**	8.404***	8.321***	0.405***
	(0.000)	(0.000)	(0.021)	(0.000)	(0.000)	(0.000)
ρ	-		0.668**	-	-	-
			(0.001)			
λ	-	0.407***	-	0.816***	-	0.333**
		(0.000)		(0.000)		(0.004)
AIC	59.724	51.024	57.33	46.159	75.293	70.776
Log Likelihood	-23.862	-19.512	-21.516	-17.079	-31.646	-29.388

	(1) W^C	(1) W^P	(1) W^R	(5) W^C	(5) W^P	(5) W^R
MORAN (Error)	3.160**	3.550***	3.234**	2.288**	2.612**	0.874
	(0.002)	(0.000)	(0.001)	(0.022)	(0.010)	(0.382)
LM Lag	6.122**	5.156**	0.571	1.018	2.131	0.936
	(0.013)	(0.023)	(0.450)	(0.313)	(0.144)	(0.333)
Robust Lag	0.243	2.774*	2.879*	1.223	2.209	6.203**
	(0.622)	(0.096)	(0.090)	(0.269)	(0.137)	(0.013)
LM Error	7.401**	2.500	5.000**	3.181*	0.734	0.085
	(0.007)	(0.114)	(0.025)	(0.075)	(0.392)	(0.770)
Robust Error	1.523	0.118	7.312**	3.386**	0.812	5.352**
	(0.217)	(0.731)	(0.007)	(0.066)	(0.367)	(0.021)

	(1)	(2)	(3)	(4)	(5)	(6)
LR Test Lag			4.691**	-		
			(0.030)			
LM Test Error			2.347	-		
			(0.126)			
LR Test Error		8.700**		13.565***		4.516**
		(0.003)		(0.000)		(0.034)
LM Test Lag		0.135		0.058		1.320
		(0.714)		(0.809)		(0.251)

Table 6: OLS and ML for Frauds, years 1999 and 2003

Independent Variables	$Fraud^{99}$			$Fraud^{99}$ W^C	$Fraud^{99}$ W^P	$Fraud^{99}$ W^R	$Fraud^{03}$		
	OLS (1)			Spatial model (2)	Spatial model (3)	Spatial model (4)	OLS (5)		
Severity	0.458*			0.409	0.520**	0.356	0.283		
	(0.093)			(0.106)	(0.042)	(0.152)	(0.295)		
Unknown	0.905***			0.813***	0.815***	0.762**	0.129		
	(0.000)			(0.000)	(0.000)	(0.001)	(0.700)		
Sind	-0.001			-0.074	-0.001	-0.022	0.232		
	(0.997)			(0.705)	(0.997)	(0.908)	(0.334)		
$Umale_{25-29}$	-0.195**			-0.122*	-0.096	-0.149**	-0.025		
	(0.005)			(0.067)	(0.160)	(0.017)	(0.822)		
$Foreigners^{t-1}$	0.181**			0.171**	0.167**	0.219**	-0.025		
	(0.012)			(0.010)	(0.012)	(0.001)	(0.714)		
Constant	4.768***			3.577***	1.724	1.982**	5.229***		
	(0.000)			(0.000)	(0.118)	(0.040)	(0.000)		
ρ	-			0.282**	0.661**	0.583**	-		
				(0.009)	(0.001)	(0.001)			
AIC	135.233			131.515	132.601	128.149	160.469		
Log Likelihood	-61.616			-58.757	-59.300	-57.074	-74.235		
	W^C	W^P	W^R				W^C	W^P	W^R
MORAN (Error)	1.573	2.325**	3.939***				0.938	0.882	1.025
	(0.116)	(0.020)	(0.000)				(0.348)	(0.378)	(0.305)
LM Lag	5.606**	4.849**	12.614***				0.381	0.036	0.483
	(0.018)	(0.028)	(0.000)				(0.537)	(0.849)	(0.487)
Robust Lag	8.144**	6.568**	5.514**				1.000	0.164	4.009**
	(0.004)	(0.010)	(0.019)				(0.317)	(0.686)	(0.045)
LM Error	1.291	0.509	7.524**				0.258	0.054	0.156
	(0.256)	(0.475)	(0.006)				(0.612)	(0.816)	(0.693)
Robust Error	4.029**	2.228	0.424				0.877	0.182	3.682*
	(0.045)	(0.136)	(0.515)				(0.349)	(0.670)	(0.055)
LR Test Lag				5.718**	4.632**	9.083**			
				(0.017)	(0.031)	(0.003)			
LM Test Error				2.741*	0.008	0.129			
				(0.098)	(0.927)	(0.719)			

Table 7: OLS and ML for Squeezes, years 1999 and 2003

Independent Variables	Squeeze^{99}			Squeeze^{99} W^C	Squeeze^{99} W^P	Squeeze^{03}		
	OLS (1)			Spatial model (2)	Spatial model (3)	OLS (3)		
Severity	0.521			0.514*	0.418	0.111		
	(0.112)			(0.095)	(0.167)	(0.680)		
Unknown	0.310**			0.278*	0.286**	0.134		
	(0.048)			(0.059)	(0.048)	(0.116)		
Sind	-0.081			-0.061	-0.074	0.167		
	(0.734)			(0.786)	(0.740)	(0.459)		
$Umale_{25\text{-}29}$	0.282***			0.198**	0.165**	0.276**		
	(0.000)			(0.012)	(0.029)	(0.011)		
$Foreigners^{t-1}$	-0.044			-0.030	-0.004	0.012		
	(0.592)			(0.700)	(0.957)	(0.859)		
Constant	1.605**			1.291*	0.557	0.707		
	(0.036)			(0.077)	(0.456)	(0.342)		
ρ	-			0.244**	0.738***	-		
				(0.032)	(0.000)			
AIC	169.336			167.297	164.887	149.433		
Log Likelihood	-78.668			-76.648	-75.444	-68.716		
	W^C	W^P	W^R			W^C	W^P	W^R
MORAN (Error)	1.083	1.961**	1.388			0.790	1.551	1.552
	(0.279)	(0.050)	(0.165)			(0.429)	(0.121)	(0.121)
LM Lag	3.986**	7.637**	1.266			0.854	1.050	1.908
	(0.046)	(0.006)	(0.261)			(0.355)	(0.306)	(0.167)
Robust Lag	10.320**	15.805***	2.385			3.908**	4.014**	6.874**
	(0.001)	(0.000)	(0.122)			(0.048)	(0.045)	(0.009)
LM Error	0.462	0.259	0.466			0.149	0.033	0.687
	(0.497)	(0.611)	(0.495)			(0.699)	(0.855)	(0.407)
Robust Error	6.796**	8.427**	1.586			3.203**	2.997*	5.654**
	(0.001)	(0.004)	(0.208)			(0.073)	(0.083)	(0.017)
LR Test Lag				4.039**	6.448**			
				(0.044)	(0.011)			
LM Test Error				6.334**	1.066			
				(0.012)	(0.302)			

5. Conclusions

In this analysis we investigated different crime phenomena, focusing on four typologies of crimes, in the Italian provinces in two years, 1999 and 2003. Interestingly the ESDA emphasised different geographies of Italian crime activities: some crimes both against persons and against property (i.e. murders and squeezes) are mostly localised nearby the organised crime (e.g. *Camorra*, *Sacra Corona Unita*, *'Ndrangheta*, *Mafia*, etc.) and other kinds of crime against property (i.e. thefts and frauds) have completely different geographies, diffused in northern provinces or all along the peninsula. Interestingly murders and squeezes follow similar patterns, while frauds and thefts some other ones, but all these phenomena changed over time showing a general dispersion over the territory.

Having this disaggregated territorial level of analysis, we were able to capture interregional differences, otherwise flattened by an analysis conducted at the regional level. Using Moran-I and Geary-C statistics to test the presence of spatial autocorrelation and using two types of spatial weights (i.e. geographical contiguity and distance), we were able to detect spatial autocorrelation affecting four kinds of crimes: provinces with high murders or squeezes are located nearby, and similarly happens to thefts and frauds, although the spatial autocorrelation is less evident.

In addition, to consider how different spillovers effects are at play, we included another weights matrix, a relational matrix based on immigration flows, to be able to capture how social networks effects influence crime activities in different provinces. Moran-I and Geary-C statistics are clear for murders and squeezes, but more ambiguous for thefts and frauds.

Once identified the existence of spatial autocorrelation, we estimated a CEM, including two types of deterrence variables (*unknown* and *severity*) and some socio-economic variables (*share of industry*, *young male unemployment* and presence of *foreigners*). Results confirm the relevance of these variables in determining crime activities, and for year 1999 they show the necessity to include a spatial term in the model, i.e. crime in province *i* is affected also by crime activities in the neighbouring regions. In 2003 these spatial effects are not captured by the model any longer, although Moran-I statistics show that presence of spatial effects.

When relational weights matrix has been used, only thefts and fraud confirm the existence of spatial autocorrelation that is captured by the error term. This should suggest that respect to less cruel crimes against property other forces are at play that are not included in this model.

Hence the use of different weights matrices to detect crime activities confirm that the geographical proximity is relevant, and that relational proximity should be appropriately investigated.

Finally respect to *severity*, this affects positively crime activities: this indicates that increasing severity with the intention to reduce the incentives to offend has not the desired effects on crime activities, on the contrary it increased crime activities, indicating the presence of a vicious circle; i.e. a person that spend lot of time in prison, once he/she served a long sentence it will be difficult to enter again in the 'society', to look for an job and to find an job because he/she does not have adequate social network knowledge and skills to change his/her living; so he/she will enter easier into local crime groups ('*peer group*') because these are 'settings' he/she know very well.

Further analyses of this model include the extension to other explanatory variables (i.e. the weight of public sector in employment) and to a spatial panel model to capture contemporarily the presence of "organised crime" phenomenon, different territorial spillovers and changing over time of a phenomenon that is also related to pardon law policies.

References

Akerlof, G.A. 1997. Social Distance and Social Decision. *Econometrica* 65: 1005–1027.

Anselin, L. 1988. *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer.

Anselin, L.2002. Under the Hood. Issues in the Specification and in the Interpretation of Spatial Regression Model. *Agricultural Economics* 27: 247–267.

Anselin, L., A.H. Bera, R. Florax and M.J. Yoon. 1996. Simple Diagnostic Test for Spatial Dependence. *Regional Science and Urban Economics* 26: 77–104.

Anselin, L. and A. Bera. 1998. Spatial Dependence in Linear Regression Models with an Application to Spatial Econometrics. In Handbook of Applied Economic Statistics edited by A. Ullah and D.E.A. Giles. Berlin/Heidelberg/New York: Springer.

Anselin, L., J. Cohen, D. Cook, W. Gorr, G. Tita. 2000. Spatial analyses of crime, in Duffee D. (ed.), *Criminal Justice*, volume 4, pp. 213-262.

Becker, G. S., (1968), "Crime and Punishment: An Economic Approach," *Journal of Political Economy,* 76, pp. 169–217.

Buonanno, P. 2006. Crime and Labour Market Opportunities in Italy (1993-2002). *Labour* 20(4): 601–624.

Buonanno, P. and L. Leone. 2006. Education and Crime: Evidence from Italian Regions. *Applied Economic Letters* 31(11): 709–713.

Burdett, K., R. Lagos and R. Wright. 2004. An on-the-Job Search Model of Crime, Inequality, and Unemployment. *International Economic Review* 45(3): 681–706.

Cezay, A.K., J. Braithwaite and B. Chapman. 1998. Unemployment and Crime: Toward Resolving the Paradox. *Journal of Quantitative Crimilogy* 14(3): 215–243.

Cherry, T.L. and J.A. List. 2002. Aggregation Bias in the Economic Model of Crime. *Economic Letters* 75: 81–86.

Dietz, R.D. 2002. The Estimation of Neighborhood Effects in the Social Sciences: An Interdisciplinary Approach. Social Science Research 31: 539–575.

Edmark, K. 2005. Unemployment and Crime: Is There a Connection?. Scandinavian Journal of Economics 107(2): 353–373.

Ehrlich, I. 1973. "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation," *Journal of Political Economy* 81, pp. 521–565.

Entorf, H. and H. Spengler. 2000. Socioeconomic and Demographic Factors of Crime in Germany. Evidence from panel data of the German States. *International Review of Law and Economics* 20: 75– 106.

Eronen, M., P. Hakola and J. Tiihonen. 1996. Factors associated with homicide recidivism in a 13-year sample of homicide offenders in Finland. *Psychiatric Services* 47: 403–406.

Florax, R.J.G.M., H. Folmer and S.J. Rey. 2003. Specification Searchers in Spatial Econometrics: The Relevance of Hendry's Methodology. *Regional Science and Urban Economics* 33: 557–579.

Glaeser, E.L., B. Sacerdote and J. A. Scheinkman. 1996. Crime and Social Interactions.*The Quarterly Journal of Economics* 111: 507–548.

Homant, R.J. 1984. Employment of Ex-Offenders: The Role of Prisonization and Self-Esteem. *Journal of Offender Counselling Services and Rehabilitation* 8(3): 5–23.

Hsieh, C.C. and M.D. Pugh. 1993. Poverity, Income Inequality, and Violent Crime: A Meta-Analysis of Recent Aggregate Data Studies. *Crime Justice Review* 18: 182–202.

Hugan, C.C., D. Laing and P. Wang. 2004. Crime and Poverty: A Search-Theoretic Approach. *International Economic Review* 45:900–938.

Ichiro, K., Kennedy B.P. R.G. Wilkinson. 1999. Crime: Social Disorganization and Relative Deprivation. *Social Science & Medicine* 48:719–731.

ISTAT. 1998. *Statistiche Territoriali*. Roma: ISTAT, (www.istat.it).

ISTAT. 1999a. *Statistiche Giudiziarie*. Roma: ISTAT, (www.istat.it).

ISTAT. 1999b. *Statistiche Demografiche*. Roma: ISTAT, (www.istat.it).
ISTAT. 1999c. *Indagine Trimestrale sulle Forze Lavoro*. Roma: ISTAT, (www.istat.it).
ISTAT. 1999d. *Dati di Contabilità Nazionale*. Roma: ISTAT, (www.istat.it).
ISTAT. 2000. *Statistiche Giudiziarie*. Roma: ISTAT, (www.istat.it).
ISTAT. 2002. *Statistiche Territoriali*. Roma: ISTAT, (www.istat.it).
ISTAT. 2003. *Statistiche Giudiziarie*. Roma: ISTAT, (www.istat.it).
ISTAT. 2003b. *Statistiche Demografiche*. Roma: ISTAT, (www.istat.it).
ISTAT. 2003c. *Indagine Trimestrale sulle Forze Lavoro*. Roma: ISTAT, (www.istat.it).
ISTAT. 2003d. *Dati di Contabilità Nazionale*. Roma: ISTAT, (www.istat.it).
ISTAT. 2006. *Statistiche Giudiziarie*. Roma: ISTAT, (www.istat.it)
Leenders, R.Th.A.J. 2002. Modeling Social Influence through Network Autocorrelation: Constructing the Weight Matrix. Social Networks 24: 21–47.
Lochner, B. 2004. Education, Work, and Crime: A Human Capital Approach. *International Economic Review* 45(3): 811–843
Luiz, J.H. 2001. Temporal Association, the Dynamics of Crime and Their Economic Determinants: A Time Series Econometric Model of South Africa. *Social Indicators Research* 53: 33–61.
Maggioni M.A., Nosvelli M., Uberti T.E. 2007. Space versus networks in the geography of innovation: A European analysis. *Papers in Regional Science*, 86: 471-493.
Marselli, R. and M. Vannini. 1997. Estimating a Crime Equation in the Presence of Organized Crime: Evidence from Italy. *International Review of Law and Economics* 17: 89–113.
Marselli, R. and M. Vannini. 2000. Quanto incide la disoccupazione sui tassi di criminalità?. *Rivista di politica economica*, Ottobre: 273–299.
Merlo, A. 2004. Introduction To Economic Models of Crime. *International Economic Review* 45: 677–679.
Messner, S.F., L. Anselin, R.D. Baller, D.F. Hawkins, G. Deane and S.E. Tolnay. 1999. The Spatial Pattering of County Homicide Rates: An Application of Explanatory Spatial Data Analysis. *Journal of Quantitative Criminology* 15(4): 423–450.
Orsagh, T. and J.R. Chen. 1988. The Effect of Time Served on Recidivism: An Interdisciplinary Theory. *Journal of Quantitative Criminology* 4(2): 155–171.
Shaw, C. R. and H.D. McKay. 1942. *Juvenile Delinquency and Urban Areas: A Study of Rates of Delinquency in Relation to Differential Characteristics of Local Communities in American Cities*, Chicago: University of Chicago.
Stigler, G. J. 1970. The Optimum Enforcement of Laws. *Journal of Political Economy* 78: 526–536.

NOTE DI LAVORO DELLA FONDAZIONE ENI ENRICO MATTEI

Fondazione Eni Enrico Mattei Working Paper Series

Our Note di Lavoro are available on the Internet at the following addresses:

http://www.feem.it/Feem/Pub/Publications/WPapers/default.htm
http://www.ssrn.com/link/feem.html
http://www.repec.org
http://agecon.lib.umn.edu
http://www.bepress.com/feem/

NOTE DI LAVORO PUBLISHED IN 2008

CCMP	1.2008	*Valentina Bosetti, Carlo Carraro and Emanuele Massetti*: Banking Permits: Economic Efficiency and Distributional Effects
CCMP	2.2008	*Ruslana Palatnik and Mordechai Shechter*: Can Climate Change Mitigation Policy Benefit the Israeli Economy? A Computable General Equilibrium Analysis
KTHC	3.2008	*Lorenzo Casaburi, Valeria Gattai and G. Alfredo Minerva*: Firms' International Status and Heterogeneity in Performance: Evidence From Italy
KTHC	4.2008	*Fabio Sabatini*: Does Social Capital Mitigate Precariousness?
SIEV	5.2008	*Wisdom Akpalu*: On the Economics of Rational Self-Medication
CCMP	6.2008	*Carlo Carraro and Alessandra Sgobbi*: Climate Change Impacts and Adaptation Strategies In Italy. An Economic Assessment
ETA	7.2008	*Elodie Rouvière and Raphaël Soubeyran*: Collective Reputation, Entry and Minimum Quality Standard
IEM	8.2008	*Cristina Cattaneo, Matteo Manera and Elisa Scarpa*: Industrial Coal Demand in China: A Provincial Analysis
IEM	9.2008	*Massimiliano Serati, Matteo Manera and Michele Plotegher*: Econometric Models for Electricity Prices: A Critical Survey
CCMP	10.2008	*Bob van der Zwaan and Reyer Gerlagh*: The Economics of Geological CO_2 Storage and Leakage
KTHC	11.2008	*Maria Francesca Cracolici and Teodora Erika Uberti*: Geographical Distribution of Crime in Italian Provinces: A Spatial Econometric Analysis

2008 SERIES

CCMP	*Climate Change Modelling and Policy* (Editor: Marzio Galeotti)
SIEV	*Sustainability Indicators and Environmental Valuation* (Editor: Anil Markandya)
NRM	*Natural Resources Management* (Editor: Carlo Giupponi)
KTHC	*Knowledge, Technology, Human Capital* (Editor: Gianmarco Ottaviano)
IEM	*International Energy Markets* (Editor: Matteo Manera)
CSRM	*Corporate Social Responsibility and Sustainable Management* (Editor: Giulio Sapelli)
PRCG	*Privatisation Regulation Corporate Governance* (Editor: Bernardo Bortolotti)
ETA	*Economic Theory and Applications* (Editor: Carlo Carraro)
CTN	*Coalition Theory Network*