UNIVERSITA' DEGLI STUDI DI BERGAMO
DIPARTIMENTO DI SCIENZE ECONOMICHE
"Hyman P. Minsky"
Via dei Caniana 2, I-24127 Bergamo, Italy
Tel. +39-035-2052501; Fax: +39-035-2052549

Quaderni di ricerca del
Dipartimento di Scienze Economiche
"Hyman P. Minsky"

Anno 2005 n.3

Criminal Activity and Education: Evidence from Italian Regions

Paolo Buonanno, Leone Leonida

- Le pubblicazioni del Dipartimento di Scienze Economiche dell'Università di Bergamo, consistenti nelle collane dei Quaderni e delle Monografie e Rapporti di Ricerca, costituiscono un servizio atto a fornire la tempestiva divulgazione di ricerche scientifiche originali, siano esse in forma definitiva o provvisoria.
- L'accesso alle collane è approvato dal Comitato di Redazione, sentito il parere di un referee.

Criminal Activity and Education: Evidence from Italian Regions

Paolo Buonanno*
University of Bergamo

Leone Leonida
University of Calabria

Abstract

This paper studies the impact of education on criminal activity in Italy. We propose a theoretical framework to determine the effects of education and past incidence of crime on criminal activity, and we test its predictions using annual data for the twenty Italian regions over the period 1980-1995. The results show that education is negatively correlated with delinquency and that crime rates display persistence over time. Our results are robust to model specifications and endogeneity.

Key words: Crime; Education; Panel Data
JEL Classification: I2; J24; K42

* *Corresponding Author.* Address: Dipartimento di Scienze Economiche, Università degli Studi di Bergamo, Via dei Caniana 2, 24127 Bergamo, Italy. *Email*: paolo.buonanno@katamail.com. We would like to thank Piergiovanna Natale, Giulio Fella, Daniel Montolio and Giovanni Caggiano for useful comments and discussions about this paper. Financial supports from the University of Milan-Bicocca and University of Calabria are gratefully acknowledged.

1 Introduction

It exists the general belief that criminals tend to be less educated than non criminals (Wilson and Herrnstein, 1985; Freeman, 1991). Prison statistics show that in 2001 more than 75% of the overall convicted population in Italy[1] had not graduated from high school and, analogously, in the US two-thirds of the incarcerated men had not attained this level of education (Freeman, 1996).

From a pure theoretical perspective, education may affect the decision to engage in criminal activities through several channels. First, higher levels of educational attainment are associated with higher return in the labour market, increasing the opportunity cost of criminal behaviour. Second, education may alter personal preferences in a way that affects decisions to engage in crime. In particular education may have a sort of "civilization" effect. Fajnzylber et al. (2002) suggest that education, incorporating a civic component, may increase the individual's moral stance, and then affect the individuals' perception of crime. Usher (1997) stresses that education perpetuates the values of society, enculturates people to serve their communities, and promotes the virtues of hard work and honesty. Furthermore, as noted by Lochner and Moretti (2004) schooling generates benefits beyond the private return received by individual. Third, school enrolment alone, independently of the level of educational attainment, reduces the time available for participating in the crime activity (Witte and Tauchen, (1994)). Finally, education also increase the cost associated with incarceration, since more educated individuals will experiences greater losses in earnings while in jail.

These theoretical implications have been empirically tested especially for the US (Lochner, 2004; Lochner and Moretti, 2004). Lochner (2004) and Lochner and Moretti (2004), using data from the National Longitudinal Survey of Youth (NLSY), found that high school graduation directly lower criminal propensity even after controlling for market returns.

[1] Data obtained by the Official Statistics of Ministero della Giustizia, year 2001.

Despite education appears to be an important variable in determining crime rates, both for its direct economic implications and for its non market effects, few papers have studied the role of education and its impact on crime.

The purpose of our paper is to study this relationship for the Italian regions. We developed a dynamic model of crime, work and education that allows us to derive some testable hypotheses on the role of education and past criminal activity on crime. Indeed, past experience in criminal activity affects in several ways the decision to commit a crime (Sah, 1991; Glaeser et al., 1996; Fajnzylber et al., 2002b); in other words, higher crime today is associated with higher crime tomorrow (i.e. persistence over time). Criminals can learn-by-doing and acquire an adequate criminal know-how level; this acquisition, in turn, makes the costs of carrying out criminal acts to decrease over time (Case and Katz, 1991). Moreover, convicted criminals have less opportunities of legal employment and a lower expected wage (Grogger, 1995). Furthermore, people may tend to have a reduced moral threshold after having joined the crime industry, to this extent the past incidence of crime in society criminal activities may lower the perceived probability of apprehension. These argument strongly suggest the possibility of criminal hysteresis or inertia.

We derive some testable hypotheses from the model we propose, which are tested by using annual data for all Italian regions over the period 1980-1995. According to general intuition and our theoretical predictions, empirical results show that education reduces delinquency, and that crime rate display persistence over time. Our results are robust to model specifications and endogeneity.

The paper proceeds as follows. Section 2 develops a time allocation model of crime, work and education, and derives our testable hypotheses. Section 3 describes our data source and the econometric methodology. Section 4 presents our empirical results. Section 5 concludes.

2 Theoretical Framework and Testable Hypotheses

Consider a representative economy where individuals decide how to allocate their available time each period to education, crime and work in the legal sector with the aim of maximizing their expected disposable income. Individuals are endowed with an initial level of ability h_0, that represents the level of ability acquired during primary school, and learning ability ε. Denote level of ability at time t by $h_t = h(s_{t-1}, \varepsilon)$, time spent for education, s_t, and time for committing crime d_t. Total time each period is normalized to 1, so time spent working is $l_t = 1 - s_t - d_t$. Each period individuals earn $w_t h_t l_t$ from legal work, where w_t is the wage rate. If an individual is engaged in criminal activity, she obtains with probability $(1 - \pi_a)$ a return $R(d_t, h_t)$, function of the time devoted to crime activities and individual ability.

Returns from crime are assumed to be strictly increasing and concave in d_t and non decreasing in h_t. With probability π_a a criminal is apprehended and punished. An apprehended criminal goes to jail for the entire period in which she is apprehended[2] and received a punishment $P(d_t)$ increasing in d_t.

The individual's maximization problem is:

$$\max_{s,d} \sum_{t=1}^{T} \beta^t y_t \tag{1}$$

under time constraints $l_t + s_t + d_t = 1$, $l_t, s_t, d_t \geq 0$, where β is the intertemporal discount rate.

Total disposable income is defined as follows

$$y_t = \begin{cases} w_t h_t l_t + R(d_t, h_t) & \text{with prob} \quad (1 - \pi_a) \\ \overline{c} - P(d_t) & \text{with prob} \quad \pi_a \end{cases} \tag{2}$$

where $\overline{c}$ the level of consumption of a convicted criminal. For simplicity we con-

[2]Our analysis is general and will not change allowing that an apprehended criminal goes to jail for a fraction of her disposable time.

sider the case in which T=2 and $s_2 = 0$.[3] Thus, we can rewrite the maximization problem in the following way:

$$\max_{s_1,d_1,d_2} \left\{ \begin{array}{l} (1-\pi_a)w_1h_1(1-s_1-d_1)+(1-\pi_a)R(d_1,h_1)+\pi_a\overline{c}-\pi_aP(d_1) \\ +\beta\left[(1-\pi_a)w_2h_2(1-d_2)+(1-\pi_a)R(d_2,h_2)+\pi_a\overline{c}-\pi_aP(d_2)\right] \end{array} \right\} \tag{3}$$

The first order conditions with respect to s_1, $d_1$and d_2 for an interior solution are:

$$d_{i,1} : (1-\pi_a)w_1h_1 = (1-\pi_a)R'(d_1,h_1) - \pi_aP'(d_1) \tag{4}$$

$$s_{i,1} : (1-\pi_a)w_1h_1 = \beta\left[(1-\pi_a)w_2h_2'(1-d_{i,2}) + (1-\pi_a)R'(d_2,h_2)h_2'\right] \tag{5}$$

$$d_{i,2} : (1-\pi_a)w_2h_2 = (1-\pi_a)R'(d_2,h_2) - \pi_aP'(d_2) \tag{6}$$

Equations (4) and (6) suggest two channels through which education can affect criminal decisions. First, education increases individual returns from work, thereby increasing the opportunity costs of crime.[4] Second, education affects the net marginal returns to crime. Equation (5) allows us to study the costs and returns of education. On the one hand, a higher education implies higher returns both from work and crime; on the other hand, an individual with a high level of education if apprehended and convicted experiences a greater earnings loss. Thus, more time invested in education in the first period is associated with higher expected returns in the legal sector in the second period, this corresponds to a higher opportunity cost of crime and then to a lower level of crime in the second period. Hence, the first testable prediction of our model is that higher level of education is correlated to less crime.

Moreover, our model suggests that the higher the fraction of time spent committing crimes in $t = 1$, the higher the fraction of time committing crime in $t = 2$.

[3] Individuals in the second period do not invest in education due that they live the last period of their life.

[4] As long as education increases the marginal return to work more than crime ($w_t h_t^{'}(s_{t-1}) > R_{h_t} h_t^{'}(s_{t-1})$), crime is decreasing in education. We argue that, for unskilled property crimes, education is likely to have a little effect on their returns.

In fact, a higher level of time spent committing crimes in the first period implies less time dedicated to schooling in the same period. This reduces expected returns in the legal sector in $t = 2$, which, in turn, lower the opportinity cost of commiting crime in $t = 2$. In this sense, our model predicts criminal inertia, which is the second testable hypothesis of our model.

Our theoretical model has other testable implications. It allows us to study the effect of wage on crime. However, we need to distinguish between first and second period. An increase of wage in the first period implies that education and crime are more costly in terms of foregone income, by choosing less education individuals will have a lower wage in the second period, then the level of crime will be higher being the opportunity cost of crime lower. On the other hand, a higher wage in the second period reduces unambiguously time spent in crime in $t = 2$. Finally, prevention and effective law enforcement policies reduce crime rate: an increase in the probability of apprehension (π_a) corresponds to a reduction of the expected return from illegal activities, and this leads to a reduction in the level of time spent in committing crime.

3 Empirical Methodology

3.1 Data

We use a panel data set for the 20 Italian regions over the period 1980-1995.[5] Table 1 describes all variables used in our estimations. Crime data are taken from CRENoS (Centre for North South Economic Research). In particular, we consider three different crime rates (*Crime*): property crime rate, theft rate and total crime rate. Crime rate is obtained normalizing the total number of recorded crime in each category by resident population in each region, population is taken from ISTAT. The explanatory variables are: education defined as average years of schooling of the population (*Education*) our own calculation on the

[5] Valle d'Aosta has been aggregated to Piemonte due to its small dimensions.

basis of the Quarterly Labour Forces (ISTAT), average regional wage ($Wage$) at 1990 constant price and percentage of crimes committed by unknown offenders ($Unknown$) as proxy of the probability of apprehension, measured as the ratio of crimes committed by unknown offenders to all recorded crimes in each category and finally GDP per capita (GDP) at 1990 constant prices.

In Table 2 we present the behaviour of crime rates per 1.000 inhabitants over the period 1980-1995. In order to facilitate the analysis we have aggregated the twenty Italian regions in northern regions, central regions and southern regions. It clearly appears that property crimes and thefts present an upward trend over the considered period. From 1980 to 1990 they have substantially remained stable, but after 1990 they peaked again. The trend is similar for all Italian regions. Total crimes present a behaviour similar to that of property crimes. Both for total crimes and property crimes the trend is overall increasing in Italy, even if over the period 1986-1990 southern regions have experienced a reduction, while central and northern regions have experienced a slight increase. Starting from 1991 a sharp increase has affected all Italian regions.

3.2 The Empirical Model and Econometric Issues

Our empirical models are as follows:

$$Crime_{i,t} = \beta_1 Crime_{i,t-1} + \beta_2 Education_{i,t} + \eta_i + \eta_t + \varepsilon_{i,t} \tag{7}$$

$$\begin{aligned} Crime_{i,t} &= \beta_1 Crime_{i,t-1} + \beta_2 Education_{i,t} + \beta_3 Wage_{i,t} + \\ &\quad \beta_5 Unknown_{i,t} + \eta_i + \eta_t + \varepsilon_{i,t} \end{aligned} \tag{8}$$

$$\begin{aligned} Crime_{i,t} &= \beta_1 Crime_{i,t-1} + \beta_2 Education_{i,t} + \beta_3 Wage_{i,t} + \\ &\quad \beta_5 Unknown_{i,t} + \beta_4 Gdp_{i,t} + \eta_i + \eta_t + \varepsilon_{i,t} \end{aligned} \tag{9}$$

Model (7) estimates a basic specification among the variables of interest. Model (8) represents a specification which is closer to our theoretical model.

Finally, model (9) adds regional GDP per capita to the regressors set, since it is a standard control variables used to capture illegal income opportunities (Ehrlich, 1973).

From an econometric perspective, there are several estimation problems that may arise in estimating these empirical models. First, we use a panel data set and it is known that OLS coefficients are biased both in the case that unobservable region-specific effects (η_i) are statistically significant and in the case that regressors and these effects (η_i) are correlated. Secondly, our model suggests that there exists a significant relationship between crime rates in t and $t-1$; for this reason, we include the lagged dependent variable in our estimating model ($crime_{i,t-1}$). In such a framework, OLS results in inconsistent estimates since $crime_{i.t-1}$ and η_i are necessarily correlated, even if the idiosyncratic component of the error term is serially uncorrelated. An obvious solution to these problems is to eliminate the term η_i by taking first-differences. However, OLS still does not consistently estimate the parameters of interest because first-differencing introduced correlation between the lagged dependent variable and differenced error terms, i.e. $crime_{i.t-1}$ and $\varepsilon_{i,t}$ are correlated trought terms $crime_{i.t-1}$ and $\varepsilon_{i,t-1}$. The alternative to first differences transformation is the within transformation; howerer altought controlling for fixed effects, the within transformation leads to consistent estimates under the hypothesis of strictly exogenous regressors.

Second, it is unlikely that explanatory variables are striclty exogenous, the relationship between crime rates and their determinants is often characterized by a two-way causality. Third, it is very likely that crime data may be subject to measurement errors, which induce biases in the estimates.

The econometric problems presented above suggest the use of an instrumental variable procedure applied to dynamic model of panel data. This paper therefore employs the GMM estimator that uses the dynamic properties of the data to generate proper instrumental variables (Arellano and Bond, 1991; Arellano and Bover, 1995). The GMM technique allows to control for (weak) endogeneity by using the instrumental variables, which consist of appropriate lagged values of the explanatory variables. To deal with the fact that measurement errors are likely to be determined not only by random errors but by specific and persistent characteristics of each region we employ the GMM-system estimator (Arellano and Bover, 1995; Blundell and Bond, 1998) which joins in a single system the regression equation in both differences and levels.

The consistency of the parameters obtained by means of the GMM estimator depends crucially on the validity of the instruments. We therefore consider two specification test suggested by Arellano and Bond (1991) and Arellano and Bover (1995). The first test is the Sargan test of overidentifying restrictions, which test the null hypothesis of overall validity of the instruments used. Failure to reject this null hypothesis gives support to the choice of the instruments. We also report the test for serial correlation of the error term, which test the null hypothesis that the differenced error term is first and second order serially correlated. Failure to reject the null hypothesis of no second-order serial correlation implies that the original error term is serially uncorrelated and the moment conditions are correctly specified.

4 Results

Table 3 shows our GMM-system estimates for property crime, theft and total crime rates. In all specifications, all variables including lagged crime rate are treated as endogenous. Four test statistics are reported: (*i*) the Wald test of joint significance of the time dummies; (*ii*) Sargan test of overidentifying restrictions; (*iii*) and (*iv*) first and second order serial correlation test.

Columns marked as (a) present the results for our basic specifications (equation 7). In all these models both education and lagged crime show expected signs. However the Sargan test rejects the null hypothesis of validity of the instruments set.

Columns (b) present the results obtained including wage rate and the proxy for the clear-up rate (*Unknown*) to the regressors set and represent the closest specifications to our theoretical model. All signs are in line with our theoritical predictions. Lagged dependent variable, the average years of schooling and the percentage of crime committed by unknown offenders have significant coefficients with the expected signs for both property crime and theft rates. Similar results are obtained for total crime rate with the exception that in this case the percentage of crime committed by unknown offenders is not significant even if has the expected sign. Finally the effect of wage on crime shows the expected sign, but it is not significant. With regards to the GMM specification tests all regression models are supported by Sargan test that confirms that the instruments used are valid (i.e. the instrument used are not correlated with the error terms). There is evidence for first-order serial correlation, while there is no evidence of second-order serial correlation; this suggests that the choice of including in the instruments set the lag $t-2$ of all variables is the correct one. Finally, time dummies are jointly significant in all the models estimated. As a further and final robustness check, in models marked as (c) we add the GDP per capita to the regressors set, as suggested by the literature. In fact, following the analysis made by Erlich (1973) we can consider the GDP per capita as a proxy for the general level of prosperity in the region, then as an indicator of illegal income opportunities. As expected GDP per capita is positively and significantly correlated to crime rate; however all other results still hold.

The econometric estimates confirm our theoretical predictions. First, crime rates display persistence over time. Second, education, measured by the average years of schooling of population, has a negative and significant effect on crime rate. With the exceptions of total crime rate, the percentage of crime committed by unknown offenders is positive and significant correlated to crime rate. This variables allows us to capture the effect of deterrence and law enforcement, a higher level of the percentage of crime committed by unknown offenders is associated to a higher expected returns from crime.

5 Conclusions

We propose a theoretical framework to determine the effects of education and past incidence of crime on criminal activity. Our model provides a set of testable hypothesis, which we test by using a GMM-system estimator for the 20 Italian regions for the period 1980-1995. Empirical results confirm our theoretical predictions. The main conclusions of this paper is that education, measured as the average years of schooling of the population, has a negative and significant effect on crime rate and that crime rates display persistence over time. These results are obtained using an instrumental variable approach that takes advantage of the dynamic properties of our data set to control for both measurement errors in crime data and joint endogeneity of the explanatory variables.

References

Arellano, M., Bover, O. (1995) Another look at the instrumental-variable estimation of error components models, *Journal of Econometrics*, 68, 29-52.

Arellano, M., Bond, S. (1991) Some tests of specification for panel data: Monte Carlo evidence and application to employment equation, *Review of Economic Studies*, 58, 277-297.

Blundell, R., Bond, S. (1998) Initial conditions and moment restrictions in dynamic panel data models, *Journal of Econometrics*, 87, 115-143.

Case, A.C., Katz, L.F. (1991) The Company You Keep: The effects of family and neighborhood on disadvantaged youth, NBER Working Paper no. 3705.

Ehrlich, I. (1973) Participation in illegitimate activities: a theoretical and empirical investigation, *Journal of Political Economy*, 81, 521-565.

Fajnzylber, P., Lederman, D. and N. Loayza (2002a), What causes violent crime?, *European Economic Review*, 46, 1323-1357.

Fajnzylber, P., Lederman, D. and N. Loayza (2002b), Inequality and violent crime, *Journal of Law and Economics*, 45, 1-40.

Freeman, R.B. (1996) Why Do So Many Young American Men Commit Crimes and What Might We Do About It?, *Journal of Economic Perspectives*, 1996, 25-42.

Freeman, R.B. (1991) Crime and the employment of disadvantaged youths, NBER Working Paper no. 3875.

Glaeser, E.L., Sacerdote, B. and J.A. Scheinkman (1996) Crime and social interactions, *Quarterly Journal of Economics*, 111, 507-548.

Grogger, J. (1995) The effect of arrest on the employment and earnings of young men, *Quarterly Journal of Economics*, 110, 51-72.

Lochner, L. (2004) Education, work and crime: a human capital approach, *International Economic Review*, 45, 811-843.

Lochner, L., Moretti, E. (2004) The effect of education on crime: evidence from prison inmates, arrests, and self-reports, *American Economic Review*, 94, 155-189.

Marselli, R., Vannini, M. (1997) Estimating a crime equation in the presence of organized crime: evidence from Italy, *International Review of Law and Economics*, 17, 89-113.

Sah, R.K. (1991) Social Osmosis and Patterns of Crime, *Journal of Political Economy*, 99, 1271-1295.

Wilson, J.Q., Herrnstein, R., 1985. Crime and Human Nature. Simon and Schuster, New York.

Witte, A.D., Tauchen, H. (1994) Work and crime: an exploration using panel data, *Public Finance*, 49, 155-167.

Table 1
Variables used in regression analysis

Variable	Description	Source
Property	Total number of property crimes recorded over population	CRENoS
Theft	Total number of thefts recorded over population	same as above
Total	Total number of crimes recorded over population	same as above
Education	Average year of schooling of the population	Author's calculation from Quarterly Labour Forces
Unknown	Percentage of crime committed by unknown offender	CRENoS
GDP	GDP per capita at constant prices 1990	ISTAT
Wage	Average wage at constant prices 1990	ISTAT

Table 2
Descriptive statistics (5 year average)

Period	North	Center	South	Italy
		Total Crimes		
1981-85	34.38	40.13	33.74	35.26
1986-90	36.89	51.09	30.75	37.36
1991-95	47.25	72.27	38.22	48.75
		Property Crimes		
1981-85	24.78	30.36	24.13	25.62
1986-90	26.20	35.48	21.44	26.23
1991-95	36.79	51.27	29.27	36.83
		Thefts		
1981-85	22.69	28.33	21.29	23.28
1986-90	23.20	32.27	17.76	22.94
1991-95	30.91	42.84	23.87	30.63

Table 3
Estimation Results: GMM-system

	Property Crimes			Thefts			Total Crimes		
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
$Crime_{-1}$	0.551	0.492	0.445	0.498	0.455	0.400	0.520	0.631	0.594
	(5.00)***	(5.56)**	(4.29)***	(4.87)***	(5.89)***	(4.44)***	(3.28)***	(4.75)***	(4.61)***
Education	-0.013	-0.006	-0.009	-0.010	-0.006	-0.009	-0.010	-0.006	-0.007
	(-1.66)*	(-2.79)***	(-2.61)***	(-1.61)*	(-2.02)**	(-2.22)**	(-1.89)*	(-1.94)*	(-1.99)**
Wage		-0.001	-0.001		-0.001	-0.001		-0.001	-0.002
		(-0.35)	(-1.08)		(-0.08)	(-1.05)		(-0.50)	(-1.46)
Unknown		0.052	0.047		0.039	0.038		0.022	0.012
		(3.04)***	(2.83)***		(1.73)*	(1.68)*		(1.37)	(1.06)
GDP			0.002			0.002			0.002
			(1.78)*			(2.11)**			(1.74)*
Wald (Time)	407.5***	267.9**	554.7**	258.5**	124.6**	334.7**	855.6**	440.5**	868.5**
Specification tests									
Sargan Test	81.88**	249.1	252.8	77.62*	245.8	249.9	174.7**	229.0	241.1
Serial Correlation									
First Order	-2.00*	-1.85*	-1.97*	-2.13*	-2.01*	-2.13*	-2.18*	-2.04*	-2.09*
Second Order	-0.45	-0.20	-0.33	0.60	0.66	0.42	-0.81	-0.74	-0.79

Notes: First Order and Second Order Test are test statistics for first and second order autocorrelations in residuals, respectively, distributed as standard normal N(0,1) under the null of no serial correlation. Sargan test is a test of overidentifying restrictions, distributed as chi-square under the null of instrument validity. Standard errors are reported in parentheses. Standard errors are robust to heteroscedasticity and autocorrelation (Arellano, 1987). ***, ** and * indicate coefficient significant at the 1%, 5% and 10% levels, respectively. All variables are instrumented using lag t-2.