Empirical Economics (2024) 66:431–465
https://doi.org/10.1007/s00181-023-02462-2

Socio-economic risk factors and wildfire crime in Italy: a quantile panel approach

Alessandra Canepa[1,2]

Received: 14 September 2022 / Accepted: 20 June 2023 / Published online: 22 July 2023

Abstract

In this paper, we analyse the socio-economic determinants of environmental crimes such as those focused on wildfire in Italy using panel data at the regional level. We also investigate the effect of economic downturns on wildfire crime. Using the nonadditive fixed effect quantile panel regression model, it was found that socio-economic factors, such as material deprivation, play an important role in driving wildfire crime. Also, risk factors such as unemployment and income inequality were seen to affect the probability of crime in the same direction. On the other hand, a negative relationship between level of education and wildfire crime was found. The results for business cycle support the conjecture that economic downturns have a significant impact on the probability of environmental crime and that the effect is particularly binding in the southern regions where unemployment and income inequality are greater. We also found evidence of a positive correlation between organised crime and wildfire crime. Once again, the grip of organised crime appears to be stronger in the southern regions.

Keywords Wildfire crime · Socio-economic factors · Quantile panel analysis

JEL Classification C21 · C33 · Q5 · Q54

1 Introduction

Wildland arson threatens lives and the stability of the environmental ecosystems since fires change the characteristics of the landscape and cause an increase in carbon dioxide emissions, accelerating climate change events (Di Fonzo et al. 2015). There are several causes of forestland fires; however, human-induced fires have been recognised as the

✉ Alessandra Canepa
Alessandra.Canepa@unito.it

1 Department of Economic and Statistics Cognetti de Martiis, University of Turin, Lungo Dora Siena 100A, Turin, Italy

2 Department of Economics and Finance, Brunel University London, Uxbridge, UK

single most important cause of forest fires (Ganteaume 2013). This is particularly the case in European Mediterranean countries where it is estimated that more than 90% of forestland fires are caused by human action (Leone et al. 2002; Velez 2009).

Against this background, the objective of this study is to investigate the impact of socio-economic factors on wildfires in Italy. The analysis was conducted using annual regional level data for the period between 2006 and 2015. The country constitutes an interesting case study for several reasons. First, in Italy, more than 10 million hectares (ha) (about 32% of the national area) were reported as natural forestland in 2022. Every year, an average of 11,000 fires occur, destroying more than 50,000 ha of wood each year according to the World Bank collection of development indicators. Italy is one of the European regions most affected by wildfires and like other Mediterranean nations (such as Portugal and Spain), and the number of wildfires has steadily increased in recent years. However, wildfires in most cases do not originate from natural causes such as lightning, spontaneous ignition, or volcanic eruptions. Instead, they are an anthropogenic phenomenon (Lovreglio et al. 2010).

In the related literature, the relationship between socio-economic factors and the incidence of crime has for a long time been an important subject of study. Theoretical models on the determinants of crime point to rational choice factors that influence the likelihood of environmental crime. For example, in the literature pioneered by Becker (1968), a criminal's choice is modelled as the standard microeconomic problem of expected utility where an individual chooses whether or not to commit a crime by comparing its expected benefits with its costs which can also include an opportunity cost, usually represented by the income from a legal activity. Ehrlich (1973) expanded on the basic analytical setting of Becker's model by introducing the interaction between potential offenders (crime supply), deterrence, and prevention (government intervention). In Ehrlich's (1973) theoretical framework, any factor that modifies the agents' opportunity cost of legal activities can be included in the analysis of the determinants of crime. The results of the empirical research support the theoretical models in that socio–economic factors play a major role in establishing the incentives for engaging in crime (Enamorado et al. 2016; Coccia 2017; Fajnzylber et al. 2002; Gould et al. 2002).

In a recent paper, Canepa and Drogo (2021) evaluated whether the socio-economic variables suggested by the general crime literature also affect wildfire crime in Italy. The authors found that social economic factors such as unemployment play an important role in driving wildfire crime (Mancini et al. 2018; Pazienza and Beraldo 2004; Michetti and Pinar 2019). The study offers several insights into the impact of socio-economic factors on the occurrence of human-caused wildfires. A possible limitation of the analysis is that the regional discrepancy of the economic background is only marginally taken into consideration. Unlike other countries in the European Union, from an economic point of view, Italy is profoundly divided. As emerges from Fig. 3 in Appendix, the GDP per capita in some regions of northern Italy is amongst the highest in the European Union, whereas most of the southern regions have the lowest per capita GDP.

In Italy, the North–South divide has been a distinctive feature of the economic development since the beginning of the twentieth century. Historically, the gap in terms of GDP growth was relatively small just after the country unification in 1861 but increased steadily over time so much so that by the 1950s, it was close to 50% of the GDP per capita in the centre–north and never significantly changed thereafter, ranging from 55 to 60% until recently. Labour productivity has also historically been significantly lower in southern Italy (Musolino 2018). Similarly, a high unemployment rate has contributed to maintaining a significant gap in income inequality and poverty rate with respect of the northern and central regions. Not only does the North–South divide constitute the main source of heterogeneity but also the differences in terms of economic development and economic performance are quite remarkable within the macro-regions. Notably, the central and northern regions have been at the core of Italian economic development since the end of Second World War, but the development model has been quite different. On one hand, the north-western regions had a development model based on the Fordist organisation, heavily relying on large firms and heavy industry, whilst on the other, the economic growth in the north-eastern regions was based on the industrial district model (A'Hearn and Venables 2013). Different models of economic development have important differences in terms of income distribution and inequality.

Against this background, in this paper, we build on Canepa and Drogo (2021) and delve further on the impact of socio-economic factors on wildfire crime focusing on regional data. In principle, if the theoretical crime models hold true (Fajnzylber et al. 2002; Ehrlich 1973), we should see a stronger relationship between the socio-economic determinants and wildfire crime in the southern Italian regions with respect of the relatively wealthier northern regions. Therefore, the question we ask in this work is the following: 'Do socio-economic factors play a greater role in driving environmental crime such as wildfire in more deprived regions?' In other words, does the North–South economic divide pave the way for wildfire crime? Also, the consensus in the crime literature agrees that the crime rate has a countercyclical behaviour, trending upward during recessions and downward during economic expansions (Bushway et al. 2019; Mehlum et al. 2006). The supporting arguments for this inverted relationship point to several, sometime contrasting, reasons. First, the quality and quantity of legitimate employment opportunities are procyclical. The higher unemployment rate associated with economic recession may promote crime by lowering the opportunity cost of the time spent in criminal activity (Grogger 1998; Gould et al. 2002; Machin and Meghir 2004). Second, the literature empirically supports that the use of criminogenic commodities such as alcohol or drug abuse are both related to the business cycle (Johansson et al. 2006; Cook and Moore 1993; Cook and Durrance 2013). Accordingly, our third research question is whether wildfire crime is also related to the economic cycle. In other words, does the countercyclical relationship found in the crime literature also hold for the type of environmental crime considered in this paper?

To answer the above questions, we conducted a two-step investigation. In the first step of the empirical analysis, and we shed light on the socio-economic drivers of wildfire crime in Italy. We also investigated these risk factors to see whether they have a higher impact on the probability of wildfire crime occurrence in poorer regions. With this target in mind, we consider a quantile regression model with a nonadditive fixed effect as suggested by Powell (2022) (see also Powell 2020). The main advantage of the procedure introduced in Powell (2022) is that the quantile estimator maintains the non-separable disturbance term traditionally associated with quantile estimation, but the instruments can be arbitrarily correlated with the nonadditive fixed effects. In general, quantile regression models are procedures used to estimate a functional relationship between the response variable and the explanatory variables for all portions of the probability distribution (Koenker 2004). Therefore, quantile regression analysis allows us to investigate whether the socio–economic risk factors considered in this work have a greater impact on lower or upper tails of the crime distribution function. Traditionally, in the empirical literature, modellers have tried to identify which environmental and socio-economic factors influence fire occurrence by using linear or nonlinear models in the context of cross-sectional data (see, for example, Martinez et al. 2009; Levi and Bestelmeyer 2016; Marchal et al. 2017; Leone, et al. 2002).[1] However, the use of pooled linear or nonlinear specifications may leave the estimated model exposed to the unobserved heterogeneity problem, leading to biased and inconsistent estimators (Cameron and Trivedi 2013). Unlike the related literature, the model adopted in this paper allows for control of the fact that the marginal effects of the socio-economic risk factors are heterogeneous throughout the wildfire crime distribution. Moreover, the adopted model is robust in terms of heteroskedasticity. In particular, using quantile panel data estimation techniques, we can account for the time-invariant region-specific factors often omitted in related empirical works. Time-invariant heterogeneity is represented by all unobservable but relevant components characterising a region which are expected to be correlated with observed factors. Examples of region-specific factors are wildfire deterrence mechanisms, land use and type, and the presence of organised crime. Whilst these instances represent a persistent problem in cross-sectional analysis, using panel data estimation techniques allow us to control for regional heterogeneity.

The remainder of this paper is organised as follows. Section 2 presents the model specification and discusses the adopted estimation methodology. Section 3 presents the results of the quantile estimation regression. Section 4 presents the effects due to economic downturns. Finally, Sect. 5 gives some conclusive remarks.

2 Model specification and methodology

In this section, we describe the model and estimation method undertaken for the empirical investigation. The proposed model specification complements and combines several approaches already developed in the related crime literature, but the estimation procedure proceeds in a different manner. Most of the empirical works investigate the relationship between wildfire crime and socio–economic factors considering how the conditional mean of wildfire crime distribution changes as a function of a given number of socio–economic risk factors (Mancini et al. 2018; Pazienza and Beraldo 2004; Michetti and Pinar 2019). To the best of our knowledge, this is the first paper that considers the causal relations and marginal effects in the tails of the wildfire crime distributions. We are particularly interested in the heterogeneity of these relationships,

[1] A noticeable exception, in this respect is the work by Michetti and Pinar (2019) where dynamic fixed effect panel data models are used to analyse the determinants of monthly variations in forest fire frequency and the size of the area burnt across Italian regions between 2000 and 2011.

in addition to the relationship of the mean and median. Below, we describe the model specification and the estimation procedure in turn.

2.1 Model specification

Let WF_{it} be the wildfire crime count in region i at time t. Consistent with the relevant literature (see Ganteaume et al. 2013 and the references therein), we assume that WF_{it} is a function of a $k \times 1$ vector of the risk factors, R_{it}, and partition R_{it} as follows:

$$R'_{it} = \left[X'_{it} Y'_{it}\right], \tag{1}$$

where the entry elements of the vector X_{it} are the covariates for social-economic risk factors, and Y_{it} is a vector of the control variables that include the demographic, environmental, and crime deterrence variables. In particular, vector X_{it} includes the risk factors related to income inequality, material deprivation, violence, educational attainments, and labour market conditions. Below, we describe the risk factors in Eq. (1) in relation to the relevant literature.

Starting with income inequality, the proxy for this covariate under consideration is the disposable income inequality, (INER_{it}). The variable $\mathrm{INER}_{\mathrm{t}}$ is defined as the ratio of the total income received by 20% of the population with the highest income to that received by 20% of the population with the lowest income. As for the expected sign of the estimated coefficients, theoretical models in the general crime literature suggest a positive relationship between income inequality and crime. For example, in his seminal paper, Becker (1968) argues that for a given probability of apprehension and expected punishment, higher levels of inequality increase the expected benefit of committing a crime for the relatively disadvantaged. In a similar theoretical framework, Fajnzylber et al. (2002) suggests that the effect of income inequality in society is strongly related to an individual's relative income. They argue that in the case of the wealthy, it is most unlikely that the rising rate in inequality induces them to commit crime. However, for poorer social actors, an increasing rate of inequality may be crime-inducing because such an increase implies a larger gap between the wages of the poor and the rich, reflecting a larger difference in income from criminal and legal activities. The idea that inequality causes crime is supported by several theoretical works (Ehrlich 1973; Imrohoroglu et al. 2000). However, strong theoretical models are not always supported by empirical evidence. The results of empirical studies reveal mixed evidence regarding the positive relation between crime rate and inequality. Whilst some have found evidence of a positive relationship between inequality and crime (Enamorado et al. 2016; Harris and Vermaak 2015; Coccia 2017; Fajnzylber et al. 2002), others have failed to find any significant relationship (Bourguignon et al. 2000; Neumayer 2005).

Closely related to income inequality is the material deprivation risk, MATDEP_{it}. In the literature, economic theory suggests that individuals are more likely to get involved in criminal activity when they experience a negative income shock. In his seminal work, Grossman (1991) established a relation between crime and material deprivation in terms of the opportunity cost framework. The author argues that decreasing income levels reduce the opportunity cost of engaging in crime with respect of other legal economic activities (Seter 2016).

Proxy variables for household wealth and the labour market conditions were also considered in the model. Namely, covariates for unemployment rate, UNEM_{it}, per capita disposable income, INC_{it}, and employment rate in non-agricultural sector, EMPL_{it}, were considered as potential risk factors. With regard to the sign of the estimated parameters, theoretical models in the general crime literature suggest a strong positive relationship between a worsening of the labour market conditions and crime. For example, Grogger (1998) estimates a structural model of time allocation between criminality, the labour market, and other non-market activities, suggesting a negative relation between wages and criminal activity. Grogger's (1998) empirical findings show that young men are responsive to wage incentives and that the racial difference in crime rate in part can attribute to the labour market. Similarly, Gould et al. (2002) found that the labour market conditions, especially wages, to be strongly related to crime for those who are most likely to commit crime (less educated men). However, concerning wildfires, the empirical studies have found mixed results. For example, Maingi and Henry (2007) found that there to be no relationship between fire occurrence and unemployment (Sebastian-Lopez et al. 2008; Martınez et al. 2009; Lovreglio et al. 2010), whereas Prestemon and Butry (2005) showed that arson fires and unemployment were related. A recent strand of the literature focuses on the indirect relation between the labour market and wildfire crime. This literature suggests that forests have been voluntarily set on fire to create firefighting jobs or to gain land for agriculture and pastures which have been retained due to being more valuable than logging (Leone et al. 2002).

To capture the effect of educational attainment on wildfire crime, two covariates were used: (i) the rate of population with an upper secondary level of education, EDUC_{it}, and (ii) the rate of population with a tertiary education, UNIV_{it}. The rationale for including the two proxy variables for education attainment is that we expected the return of education on income to be higher for individuals with a university degree.

With respect of the causal relationship between crime and educational attainment, in the literature, it has been argued that an individual's education level may impact their decision to commit a crime in several ways. First, the effect of income is positively related to education. This is because higher levels of education attainment can be associated with increasing returns of legitimate work and a rise in the opportunity costs of illegal behaviour (Lochner 2004; Lochner and Moretti 2004). Second, the resources allocated to education create time constraints that deter criminal offences. Tauchen et al. (1994) investigated this "self-incapacitation effect" and found that the time spent on pursuing education is negatively correlated with the probability conviction amongst youngsters. Hjalmarsson (2008) focused on the impact of getting arrested before finishing school on the probability of graduating from high school and found that the probability of a young person being convicted for committing a crime greatly increases their likelihood of becoming a high-school dropout. Third, a stream of literature also associates a greater education with higher life satisfaction which, in turn, reduces the probability of committing a crime. For example, Oreopoulos (2006) and Lochner (2004) suggest that higher levels of education can increase risk aversion, lowering the crime rate. In an interesting paper, Usher (1997) argues that education promotes a "civilisation effect" that contributes to reducing the incidence of criminal activity. The author argues that education conveys a civic externality, a benefit to society over and above the benefit to the student in terms of enhancing his future earning power. Michetti et al. (2019) analysed the determinants of the monthly variations in forest fires for Italian regions between 2000 and 2011 and concluded that education attainment plays an important role in preventing fraudulent activity. Similarly, Torres et al. (2012) found that the areas with fewer fires are characterised by a population with higher levels of education.

The last risk factor considered is the level of violence. Also, in this case, two proxies for violence were considered: (i) the homicide rate (HOMR_{it}) and (ii) organised crime, (ORGC_{it}), defined as the conviction rate for organised and mafia-related crime. The rationale for including the homicide rate as well as organised crime convictions is that the latter variable likely suffers from a significant measurement error. Organised crime is a difficult phenomenon to capture and using the number of trials for organised crime as the sole covariate to assess the impact of organised crime on wildfire crime may not be informative. In this respect, the inclusion of the covariate "homicide rate" may be useful to signal the significant presence of organised crime in a region. Clearly, the overall homicide rate does not distinguish between homicides committed by criminal organisations and other homicides. On the other hand, it is unlikely to suffer from a measurement error and allows us to test the hypothesis that the degree of violence in a region has an effect on the occurrence of wildfire crime.

As far as the sign of the expected estimated coefficients is concerned, in the literature, studies on the relationship between organised crime and wildfire crime are rare. One of the few empirical works that considers this relationship is the EFFACE (2016) report where the evidence found there to be a positive relation between organised crime (mafia-like organisations) and the rate of fire crimes. The influence of organised crime is reported to be stronger in Italy's southern regions where the government's ability to enforce the law is weaker. The literature on environmental crime mainly focuses on the growing role of organised crime in relation to other types of environmental crime. This is particularly the case for illegal dumping and the international illegal trafficking of hazardous waste where it was found that organised mafia-like criminals play a significant role in environmental criminality (Germani et al. 2018). Overall, being the socio-economic determinant of other types of environmental crime similar to wildfire crime, we expect a positive relationship between organised crime and arson.

Regarding the control variables, the entry elements of vector Y_{it} in Eq. (1) are (i) deterrence factors, (ii) weather-related factors, and (iii) demographic risk factors:

$$Y_{it} = [\mathrm{TRIAL}_{it}, \mathrm{EQ}i_{it},\ \mathrm{RAIN}_{it}, \mathrm{TEMP}_{it},\ \mathrm{AGRI}_{it}, DEN_{it}]. \tag{2}$$

In Eq. (2), the covariate TRIAL_{it} is a proxy that captures the probability of apprehension. According to the literature, a higher deterrence level reduces the level of wildfire crime (Canepa and Drogo 2021 and the references therein), and we therefore expect a negative estimated sign between WF_{it} and TRIAL_{it}.

The variable $\mathrm{EQ}i_{it}$ is a proxy for the quality of institutions. The quality of governance and institutions is another important deterrence factor which not only affects wildfire crime directly but also affects the socio-economic outcomes such as education, poverty, and inequality; hence, it has an indirect effect on forest fires through its effect on these socio-economic factors.[2] We therefore expect there to be a negative correlation between this covariate and wildfires (Charron et al. 2019).

Concerning weather-related factors, in Eq. (2), the variable RAIN_{it} is the annual precipitation in mm and TEMP_{it} is the temperature, measured as the mean temperature. In general, weather conditions that cause downward changes in fuel moisture and, consequently, upward changes in fuel availability are expected to increase the probability of wildfire occurrence (Albertson et al. 2009; Plucinski 2014; Guo et al. 2016). Similarly, higher mean and maximum temperatures are expected to exhibit a positive relation with wildfires (Preisler et al. 2004; Carvalho et al. 2008).

Looking at the population density, DEN_{it}, in the fire related literature, an increase in population density has been found to be positively related to wildfire crime. For example, Catry et al. (2007) observed that a large majority of the fire ignitions in Portugal occurred in the municipalities with the highest population densities. Gonzalez-Olabarria et al. (2015) found that the distribution of arson in north-eastern Spain occurred near coastal areas where the population density was higher. Similarly, Romero-Calcerrada et al. (2008) found there to be a positive relationship between the intensive use of the territory and ignitions in the forest areas in Spain (Padilla and Vega-Garcıa 2011).

Finally, AGRI_{it} is the proportion of population in agricultural employment. Socio-economic transformations in rural areas such as rural exodus, a reduction in agricultural employment, and the abandonment of agricultural land may contribute to wildfire crime. In the related literature, the impact of this covariate is controversial with several empirical studies reporting a positive relation between fire occurrence and agricultural activities (Martınez et al. 2009; Rodrigues et al. 2016 amongst others), whilst others do not find that a decreased human impact associated with agricultural land abandonment leads to a statistically significant decrease in fire ignition probability (Ricotta et al. 2012).

2.2 Estimation methods

To investigate the effect of the risk factors in Eq. (1), we considered a quantile regression model with a nonadditive fixed effect as suggested by Powell (2022) (see also Chernozhukov and Hansen 2008; Powell 2020). The adopted model specification is particularly convenient as it allows us to shed some light on two related questions. First, is there a causal relation between the socio-economic risk factors and wildfire crime? Second, does one unit increase in a given risk factor of vector R_{it} in Eq. (1) affect the regions with a lower wildfire crime rate differently from the regions with a higher wildfire crime rate?

To introduce the model under consideration, we relate it to the linear regression specification that is usually adopted in the literature. Most of the empirical works

[2] We are grateful to two anonymous Referees for suggesting these covariates to us.

consider the conditional distribution of $E(\mathrm{WF}|R)$ using a structural model:

$$\mathrm{WF} = \alpha + R'\beta + \varepsilon, \tag{3}$$

where α and β are unknown constant parameters, and ε is an error term.[3] In Eq. (3), the parameter β has a causal interpretation as some effect of the risk factor R on WF. However, for the Italian case, assuming that the marginal effects of the risk factors are the same in all regions may not be informative due to the heterogeneity of the regional economy. Rather than letting the differences in marginal effects go into the error term ε and interpret β as some sort of average, we try to learn about the marginal effect heterogeneity.

In this paper, we allow the parameter β to be a "random coefficient", and in other words, the structural model we consider allows for region-specific parameters. Since in our model, β is a vector of region-specific coefficients, they are no longer constants like in Eq. (3), but rather random variables in a "random coefficients" model. The additive error term ε in Eq. (3) is now redundant since it is absorbed into the random intercept.

To make the model more tractable, we follow Powell (2022) and assume that β is a deterministic (but unknown) vector-valued function $\beta(\cdot)$ applied to an unobserved random variable U:

$$WF_{it} = R_{it}^{'}\beta_j\left(U_{it}^{*}\right), \tag{4}$$

where $U_{it}^{*} = f(\alpha_i, \varepsilon_{it})$ for some fixed effect α_i, and unknown function f. Assuming that U_{it}^{*} is continuous, it can be normalised to have a Unif(0, 1) distribution. The model in Eq. (4) allows the parameters to vary based on an unspecified function of the fixed effect and an observation-specific disturbance term whilst permitting individual-specific heterogeneity. Under the monotonicity assumption, $R_{it}^{'}\beta_j(U_{it}^{*})$ is an increasing function of U, then

$$P\left[\mathrm{WF}_{it} \le R_{it}^{'}\beta(\tau)|Z_{it}\right] = \tau,$$

where $\tau \in (0, 1)$ and $Z_{it} = \left(Z_{i1}, \ldots, Z_{i,T}\right)$ is a set of instruments that can be arbitrarily correlated with the nonadditive fixed effect.

The model in Eq. (4) can be estimated using the GMM method in an instrumental variable context under the following moment condition

$$E\left\{(Z_{it} - Z_{is})\left[1\left(WF_{it} \le R_{it}^{'}\beta(\tau)\right) - 1\left(WF_{is} \le R_{is}^{'}\beta(\tau)\right)\right]\right\} = 0, \text{ for all } s, t,$$

$$E\left[1\left(WF_{it} \le R_{it}^{'}\beta(\tau)\right) - \tau\right] = 0. \tag{5}$$

[3] Note that, to simplify the notation, we refer the general regression model in Eq. (3) to the vector of risk factors in Eq. (1) without loss of generality.

Under condition in Eq. (5), in addition to other standard regularity conditions, the GMM estimation method produces consistent and asymptotically normal estimators (for more details, see Powell 2022).

2.2.1 Model specification issues

In the model specification procedure, two issues were taken into consideration. First, in Eq. (1), more than one proxy variable for a given socio-economic risk factor was considered and some of them may be highly correlated. Accordingly, a number of models were estimated using the subset of the variables included in Eq. (1). In the model selection procedure, the possible presence of multicollinearity was assessed using the eigenvalues of the different covariance matrices and computing the conditioning index (CI). Following Pena and Renegar (2000), the CI index was calculated as

$$\mathrm{CI} = \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}} \tag{6}$$

where $\lambda_{\max}$ is the largest eigenvalue and $\lambda_{\min}$ the smallest eigenvalue of the variance covariance of the risk factor matrix R_{it}. If CI $= 1$, there is no evidence of collinearity between the covariates in the estimated model. However, as the collinearity increases, the eigenvalues in Eq. (6) become both greater and smaller than 1 and the CI number increases. Pena and Renegar (2000) suggest that CI < 10 for a well-defined matrix as opposed to $11 <$ CI < 30 for moderate multicollinearity. Accordingly, the condition we imposed for introducing a covariate to the final model specification was a calculated CI number that was less than 10.

The adopted model specification procedure resulted in the selection of the best fitting models from a set of candidate specifications. We labelled these models M1, M2, M3, and M4, respectively. Models M1 and M2 include MATDEP_{it} as the risk factor, whereas, in models M3 and M4, the risk factor UNEM_{it} is included. However, the covariate MATDEP_{it} is excluded. The reason for this specification is that these two covariates were found to be highly correlated and, therefore, only one risk factor at the time was considered. Similarly, the two pairs of covariates of EDUC_{it} and UNIV_{it}, and INC_{it} and EMPL_{it} produced a high CI index. Accordingly, the EDUC_{it} risk factor was included in models M1 and M2 but excluded from models M3 and M4 where the covariate was replaced by UNIV_{it}. To avoid multicollinearity, INC_{it} was included in model M1 but not in model M2, and it was excluded from the model specification when the risk factor UNEM_{it} was included as the covariate (models M3 and M4). Surprisingly, the two proxies for violence, ORGC_{it} and HOMR_{it}, were not found to be highly correlated, and the CI index in all estimated models was found to be less than 10. For this reason, both risk factors were included in the models. The condition index for the estimated models is presented in Table 6 along with the correlation matrix of the covariates under consideration (Table 5).

The second issue in relation to the estimation of Eq. (4) is the choice of instrumental variables for the potentially endogenous covariates. The related literature suggests that the inequality variable may be endogenous since income inequality may be correlated with crime (see Enamorado et al. 2016; Harris and Vermaak 2015; Fajnzylber et al. 2002). However, the reverse may also be true. To mitigate concerns about this form of reverse causality, we constructed an instrumental variable that is correlated with changes in regional inequality but not associated with regional wildfire crime rates. Specifically, we followed Enamorado et al. (2016) (see also Bartik 1991; Blanchard and Katz 1992; Boustan et al. 2013) and constructed the instrument by calculating the predicted inequality growth rate by interacting the regional inequality shares with the national inequality growth rate. In particular, for each year, we estimated the percentile share of inequality. In doing so, we estimated to which national percentile of inequality each region belonged to in the initial year. We then used the estimated lagged shares of inequality and interacted them with the change in inequality at the national level. By design, this instrument allowed us to isolate the changes in inequality at the regional level that is driven by national shifts which should be correlated with regional welfare indicators but not with the wildfire crime observed in each region. In addition to this instrument, we followed Aldieri and Vinci (2017) (see also Powell 2022) and used lagged explanatory variables as instruments. The idea is that the lagged level of the covariates may impact the trends in risk factors but do not share a statistically significant relationship with the current wildfire crime rate. This should also control for the potential endogeneity of the other covariates.[4]

3 Data and descriptive statistics

For the empirical investigation, we used the annual data for the period 2006–2016 for twenty Italian regions. Table 1 provides a description of the acronyms and some details of the authors' calculations.

Before reporting the estimation results, we undertook a preliminary descriptive statistical analysis of the quantile distribution of the variables under consideration in order to investigate the properties of the cumulate distribution functions of wildfire crime and the related socio-economic determinants. Table 2 reports the estimates of the quantile shares of the socio-economic risk factors along with wildfire crime.[5]In particular, the first column reports the risk factors, whereas in columns 2–11, the estimated proportion of the total outcome that falls into each quantile is described. For

[4] Please note that in the general crime literature, the crime deterrence variables are often found endogenous (see for example Eck et al. 2000). Accordingly, endogeneity tests were carried to check for the potential endogeneity of the covariate TRIAL_{it}. However, we failed to reject the null hypothesis of endogeneity, we therefore concluded that using the lagged covariate as a precautionary instrument was enough to resolve the issue of endogeneity. Similar results were also found in Canepa and Drogo (2021). Endogeneity tests were also carried out for UNEM_{it} and MATDEP_{it}, however, also in this case, the tests statistics fail to reject the null hypothesis of endogeneity of these covariates. The results are not reported, but available upon request.

[5] To calculate the quantile share, we consider the distribution function of each element of the vector.

$$\widetilde{X} = [\text{WF}, \text{INER}, \text{ORGC}, \text{HOMR}, \text{EDU}, \text{MATDEP}, \text{UNEM}, \text{INC}].$$

Let j (for $j = 1, \ldots, 8$)*be an element of the vector* $\widetilde{X}$, then, for any j, we define $F_j(\widetilde{X}) = \Pr\{\widetilde{X} \leq x\}$ as the distribution function of X_j and the quantile function as.

$$Q_j(p) = F^{-1}(p) = \inf\{x|F(x) \geq p\},$$

Table 1 Variable list and descriptive statistics *Source* The Italian National Institute for Statistics (ISTAT), the QoG Institute (University of Gothenburg) and authors' calculations. Note that for the EQI index the missing observations were obtained using linear interpolation between 2010 and 2013

Variable Name	Description	Mean	Standard Deviation
WF_{it}	Arson crime rate (per 100,000 persons). The variable includes the number of reported cases of deliberate wildfire with known offender plus the total cases of unknown offender. Unintentional forest fires are not included	10.63	13.37
$INER_{it}$	Quintile share ratio (S80/S20) for disposable income: ratio between average income of the top quintile and average income of the bottom quintile. The variable is calculated using an income equivalent factor to account for the heterogeneity of family compositions, such as the different needs between children and adults, for example, or economies of the scale generated by sharing the same dwelling[a]	5.04	1.08
$ORGC_{it}$	Organised crime and mafia-related crime rate (per 100,000 persons)	2.07	1.51
$HOMR_{it}$	Homicides rate per 100,000 inhabitants	0.84	0.59
$EDUC_{it}$	Percentage of the population aged 25–64 with secondary education attainment	41.12	4.71
$UNIV_{it}$	Percentage of the population aged 25–64 with tertiary education attainment	15.26	2.67
$EMPL_{it}$	Employment in non-agricultural sector for the working-age population	825,458.7	802,060.3
INC_{it}	Per Capita Income	17,462.4	3445.02
$MATDEP_{it}$	Principal component analysis of: *i*) proportion of household in economic distress, and *ii*) proportion of household leaving in severe material deprivation. (see Table 5 for details)[b]	9.75	1.00

Table 1 (continued)

Variable Name	Description	Mean	Standard Deviation
$UNEM_{it}$	Total unemployment rate	9.70	4.95
$TRIAL_{it}$	Share of forest fire related cases that go to trial out of the cases with an identified suspect	0.610	0.228
EQI_{it}	European Quality of Government Index (EQI) is based on large citizen surveys in 2010 and 2013 where respondents are asked about perceptions and experiences with public sector corruption, along with the extent to which citizens believe various public sector services are impartially allocated and of good quality	3.95	39.06
DEN_{it}	Population density: total population in a region divided by the total area in square kilometres	183.99	111.63
$RAIN_{it}$	Total precipitation in mm per year	826.67	201.88
$TEMP_{it}$	Average temperature in Celsius in a year	13.43	0.29
$AGRI_{it}$	Percentage of employment rate in the agricultural sector	4.57	2.49

[a]Note that the Gini coefficient has often been used in the literature to investigate the relationship between crime and income inequality (see for example Fajnzylber et al. 2002); however, this index may be biased towards the central part of the income distribution. For this reason, the income quantile share was considered as a proxy for income inequality

[b]two proxy variables for low income were considered: (i) proportion of household in economic distress and (ii) the proportion of household leaving in severe material deprivation. However, due to the high correlation between these two variables (the calculated correlation coefficient is above 80%), and with the aim to reduce the number of regressors in estimation, we used the technique of principal component analysis over these two indicators to construct a composite measure of material deprivation risk. The resulting risk factor is referred to as $MATDPR_{it}$. Details of the calculation are reported in Appendix

Footnote 5 continued

with $p \in [0, 1]$. For each element of $\widetilde{X}$ let the proportion of the total outcome that falls into the quantile interval $\left(Q_{j,p_{l-1}},\ Q_{j,p_l}\right)$, for $p_{l-1} \leq p_l$ be

$$S_j(p_{j,l-1}, p_{j,l}) = \frac{\int_{-\infty}^{Q_{p_l}} x\mathrm{d}F(x) - \int_{-\infty}^{Q_{p_{l-1}}} x\mathrm{d}F(x)}{\int_{-\infty}^{\infty} x\mathrm{d}F(x)}. \tag{7}$$

The expression in Eq. (7) defines the quantile shares and allows us to investigate the shares of total outcome pertaining to the population segment from relative rank $Q_{j,p_{l-1}}$ to relative rank Q_{j,p_l} in the list of the ordered outcomes. Note that, to simplify the notation, the i and t subscripts for each variable have been dropped below.

Table 2 Estimated proportions of total outcome by percentile shares in percentage

Risk factors	Quantiles									
	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
WF	0.73 (0.110)	1.54 (0.183)	2.25 (0.249)	3.30 (0.377)	4.89 (0.469)	6.57 (0.538)	9.43 (0.689)	12.29 (0.707)	17.22 (0.817)	41.73 (2.782)
INER	7.66 (0.112)	8.18 (0.111)	8.58 (0.112)	8.90 (0.109)	9.21 (0.101)	9.56 (0.127)	10.12 (0.111)	10.86 (0.127)	12.03 (0.195)	14.85 (0.474)
ORGC	2.39 (0.474)	4.88 (0.255)	5.75 (0.297)	6.79 (0.307)	7.75 (0.305)	8.88 (0.348)	10.55 (0.437)	12.40 (0.452)	15.44 (0.551)	25.14 (2.140)
HOMR	1.89 (0.427)	4.32 (0.504)	6.05 (0.319)	7.13 (0.340)	7.67 (0.331)	9.13 (0.344)	10.48 (0.362)	12.10 (0.445)	14.69 (0.471)	26.52 (1.764)
EDUC	7.93 (0.064)	8.52 (0.087)	9.12 (0.123)	9.64 (0.086)	10.03 (0.047)	10.34 (0.057)	10.64 (0.055)	10.88 (0.064)	11.17 (0.070)	11.67 (0.102)
MATDEP	2.89 (0.141)	3.84 (0.200)	4.79 (0.210)	5.74 (0.208)	6.96 (0.307)	8.62 (0.359)	11.89 (0.735)	15.07 (0.455)	18.09 (0.484)	22.08 (0.822)
UNEM	3.48 (0.171)	4.92 (0.216)	5.95 (0.266)	7.24 (0.249)	8.43 (0.242)	9.75 (0.263)	11.24 (0.255)	12.84 (0.250)	14.90 (0.564)	21.19 (0.596)

The table reports the quantile share distributions for the socio-economic risk factors. The quantile shares were calculated as explained in footnote 5 ease of interpretation, the proportions are reported as a percentage and the quantiles are expressed in terms of percentile shares.

From Table 2, it appears that wildfire crime density function is highly skewed since approximately 60% of wildfire crimes are concentrated in the regions located in the higher quantiles of the cumulate distribution function. Namely, the regions in the top 30th quantile account for approximately 60% of the proportion of wildfire crime in the sample under consideration. On the other hand, the regions in the bottom 30th quantile experience less than 5% of wildfire crime. It is interesting to note that the distributions of organised crime and homicide rate, along with most of the other socio-economic indicators, seem to follow a similar pattern. In particular, Table 2 shows that the regions in the top 30th quantiles contribute to 52.98% of organised crime convictions and 53.31% of homicides out of the total crimes, respectively. This contrasts with the regions in the bottom 30th quantile where the percentile shares decrease to 13.02% for organised crime convictions and to 12.26% for homicide prosecutions, respectively. Similarly, the distribution of poverty and unemployment is highly skewed with the top 30% worst-performing regions in the sample receiving 55.24% and 48.93% of the total share in terms of poverty and unemployment, respectively. Looking at inequality, the ratio between the average income of the top quintile and the average income of the bottom quintile is also a left-tailed distribution with 37.74% of the share being in the higher top 30$^{\text{th}}$ quantiles. Finally, the education distribution seems to be quite uniform across the Italian regions with only marginal differences between the top- and worst-performing regions.

4 Empirical results

In Table 3, the estimated coefficients of the four best fitting models M1, M2, M3, and M4 are reported. For brevity, only the estimated parameters for the quantiles $p \in \{0.25, 0.50, 0.75\}$, along with the associated standard errors, are reported. Note that most of the covariates are considered in natural logs.

In relation to our first question in Sect. 2.2, looking at the estimating results for the median quantile in Table 3 reveals interesting insights on the causal relation between socio-economic risk factors and wildfire crime. In particular, from the middle panel of Table 3, it appears that most of the socio-economic risk factors are significantly different from zero and have the expected sign. In particular, it appears that there is a positive relation between income inequality and wildfire crime as the estimated signs for INER_{it} in models M1–M4 are uniformly positive, and the estimated coefficients are statistically significant. Following a theoretical argument in Becker (1968), the positive effect of income inequality on wildfire crime can be interpreted according to the cost–benefit framework where the magnitude of the income inequality coefficient measures the impact of the difference between the return to crime and its opportunity cost (measured by the legal income of poor individuals). However, this argument is purely based on the rational behaviour of agents which, in the case of environmental crime, may underestimate the importance of the socio–economic environment in which the social actors interact. Following Fajnzylber (2002), an alternative interpretation of the positive link between inequality and crime is that in regions with higher

Table 3 Risk factors of wildfire crime rate

Quantiles	0.25				0.50				0.75			
Risk factors	M1	M2	M3	M4	M1	M2	M3	M4	M1	M2	M3	M4
	Coef	Coef	Coef	Coef	Coef	Coef	Coef	Coef	Coef	Coef	Coef	Coef
$\ln\text{INER}_{it}$	0.055^{**} (0.027)	0.081^{***} (0.007)	0.167^{***} (0.042)	0.600^{***} (0.107)	0.080^{***} (0.013)	0.218^{***} (0.039)	0.670^{***} (0.100)	1.431^{***} (0.006)	0.557 (0.128)	2.231^{***} (0.169)	1.915^{***} (0.050)	1.619^{***} (0.026)
MATDEP_{it}	1.365^{***} (0.035)	0.762^{***} (0.013)	–	–	1.432^{***} (0.176)	1.681^{***} (0.036)	–	–	3.668^{***} (0.109)	5.935^{***} (0.029)	–	–
$\ln\text{UNEM}_{it}$	–	–	1.851^{***} (0.103)	1.209^{***} (0.352)	–	–	1.888^{***} (0.247)	1.288^{***} (0.010)	–	–	1.978^{***} (0.070)	2.118^{**} (0.020)
$\ln\text{EDUC}_{it}$	0.010 (0.013)	– 0.035^{***} (0.004)	–	–	– 0.095^{***} (0.034)	– 0.077^{***} (0.025)	–	–	– 0.753^{***} (0.014)	– 0.299^{***} (0.039)	–	–
$\ln\text{UNIV}_{it}$	–	–	– 0.064^{***} (0.023)	– 0.934^{***} (0.026)	–	–	– 0.227^{***} (0.033)	– 1.115^{***} (0.027)	–	–	– 0.882^{***} (0.091)	– 0.673^{***} (0.008)
$\ln\text{ORGC}_{it}$	0.770^{***} (0.018)	0.331^{***} (0.005)	0.716^{***} (0.014)	3.109^{***} (0.096)	0.687^{***} (0.098)	0.247^{***} (0.051)	0.505^{***} (0.038)	2.708^{***} (0.012)	0.294^{***} (0.048)	– 0.813^{***} (0.143)	0.398^{***} (0.020)	3.841^{***} (0.028)
$\ln\text{HOMR}_{it}$	3.068^{***} (0.028)	4.416^{***} (0.013)	4.651^{***} (0.097)	–	4.855^{***} (0.171)	5.701^{***} (0.215)	5.289^{***} (0.062)	–	6.248^{***} (0.091)	8.003^{***} (0.326)	15.848^{***} (0.142)	–
$\ln\text{INC}_{it}$	2.588 (0.430)	–	–	–	– 1.119 (0.407)	–	–	–	- 15.256 (0.242)	–	–	–

Table 3 (continued)

Quantiles	0.25				0.50				0.75			
Risk factors	M1 Coef	M2 Coef	M3 Coef	M4 Coef	M1 Coef	M2 Coef	M3 Coef	M4 Coef	M1 Coef	M2 Coef	M3 Coef	M4 Coef
$\ln EMPL_{it}$	–	– 1.818 (0.017)	–	–	–	– 2.212*** (0.074)	–	–	–	– 6.815*** (0.298)	–	–
$\ln TRIAL_{it}$	– 2.646*** (0.092)	– 2.342*** (0.015)	– 2.046*** (0.061)	– 1.165*** (0.131)	– 3.521*** (0.195)	– 1.947*** (0.149)	– 3.327*** (0.088)	– 1.127*** (0.007)	– 3.990*** (0.093)	– 2.699*** (0.396)	– 3.627*** (0.094)	– 2.575*** (0.036)
$\ln EQI_{it}$	–	–	–	– 0.024*** (0.000)	–	–	–	– 0.047*** (0.000)	–	–	–	– 0.084*** (0.000)
$\ln DEN_{it}$	– 0.012*** (0.000)	– 0.002*** (0.000)	– 0.015*** (0.000)	– 0.034*** (0.001)	– 0.019*** (0.000)	– 0.005*** (0.000)	– 0.027*** (0.000)	– 0.018*** (0.000)	– 0.003*** (0.000)	– 0.012*** (0.003)	– 0.036*** (0.004)	– 0.026*** (0.000)
$\ln RAIN_{it}$	– 0.004*** (0.000)	– 0.003*** (0.000)	– 0.001*** (0.000)	– 0.008*** (0.000)	– 0.001*** (0.000)	– 0.003*** (0.000)	– 0.003*** (0.000)	– 0.007*** (0.000)	– 0.003*** (0.000)	– 0.008*** (0.001)	– 0.001*** (0.000)	- 0.004*** (0.000)
$\ln TEMP_{it}$	0.539*** (0.020)	0.277*** (0.018)	0.455*** (0.072)	0.143* (0.077)	0.301*** (0.045)	0.401*** (0.029)	0.537*** (0.022)	0.599*** (0.002)	0.326*** (0.013)	0.094 (0.065)	0.437*** (0.051)	0.746*** (0.012)
$\ln AGRI_{it}$	0.577*** (0.081)	– 0.118*** (0.510)	– 0.344*** (0.078)	– 1.258*** (0.764)	0.121 (0.510)	– 1.347*** (0.199)	−2.563*** (0.256)	– 1.271*** (0.007)	7.875*** (0.127)	3.661*** (0.396)	7.396*** (0.094)	2.575*** (0.036)

Table 3 (continued)

Quantiles	0.25				0.50				0.75			
Risk factors	M1 Coef	M2 Coef	M3 Coef	M4 Coef	M1 Coef	M2 Coef	M3 Coef	M4 Coef	M1 Coef	M2 Coef	M3 Coef	M4 Coef
Number of observation	167	144	152	48	167	144	152	48	167	144	152	48
Number of years	9	9	9	3	9	9	9	3	9	9	9	3

The table report the estimated coefficients of the GMM-instrumental variable quantile panel estimations with dependent variable WF_{it}. The related standard errors are presented in parentheses below their corresponding coefficients

***, **, *denote statistical significance at 1%; 5%; and 10%, respectively

Instruments: lagged explanatory variables for all the covariates in the model in addition to the interaction share inequality variable calculated as explained in Sect. 2.2. The "ln" indicates that the covariates were taken in natural logs income inequality, individuals have low expectations of the lifetime improvement of their socio-economic status through legal economic activities which decreases the opportunity cost of participating in illegal endeavours more generally.

Regarding the effect of violence, the estimated coefficients for $ORGC_{it}$ in the median quantile are positive and statistically significant throughout the estimated models, whereas the estimated parameters for $HOMR_{it}$ are of the expected sign but significant only in models M1, M2, and M3. These results are important since, according to a report by the Italian Antimafia Directorate, environmental crime is currently one of the most profitable forms of criminal activity, and eco-mafia has become a big business in the waste sector in Italy (Savona and Riccardi 2018). The previous literature has found that organised criminal networks are involved in the illegal disposal of commercial, industrial, and radioactive waste (Germani et al. 2018). However, the relation between organised and wildfire crime has not been previously investigated. Bearing in mind that organised crime is deeply rooted in Italy, and this result may have important implications when it comes to the crime deterrence policy.

Focusing on the role of education, the results for the median quantile in Table 3 suggest that a negative relationship exists between education attainment and wildfire crime throughout the estimated models. A negative relation between education level and fire risk was also found to be significant by Torres et al. (2012). Regarding unemployment, the estimated parameters for the median quantile in models M3 and M4 support the view that the worsening of the labour market conditions increases wildfire crime. In general, the findings in Table 3 confirm the previous empirical results in the related literature that unemployment is an important risk factor causing wildfire crime (Mercer and Prestemon 2005; Prestemon and Butry 2005; Martınez et al. 2009).

Coming now to $TRIAL_{it}$, and EQI_{it}, we see that higher enforcement levels and institution quality reduce the probability of committing a crime since the estimated coefficients of these covariates are negative throughout the estimated the models, as postulated by the seminal paper by Becker (1968). In particular, we see that the wildfire crime rate is negatively related to the share of suspected wildfire criminals who have had to stand trial. This implies that the successful identification of suspected offenders increases the expected costs of wildfire crime for the average potential arsonist. Overall, the estimation results strongly suggest that a higher level of apprehension and good governance helps to reduce wildfire crime.

Focusing on the demographic covariates, the estimated coefficients of DEN_{it} for the median quantile have a negative sign and are statistically significant. However, the estimated parameters in the models are very small. Such a small magnitude of estimated coefficients casts some doubts on the actual impact of this covariate on wildfire crime for the data at hand.

Coming to weather-related factors, the results mainly confirm the findings in the related literature as the estimated parameters in Table 3 are significant and of the expected sign. It is well known that the Mediterranean summer weather conditions (high temperature, prolonged drought periods, and strong winds) facilitate wildfire occurrence (Vasilakos et al. 2009).

Turning now to the second question in Sect. 2.2, comparing the estimated coefficients of the lower and top quantiles reveal some interesting results. Starting with income inequality, it appears that the magnitude of the estimated coefficients for $INER_{it}$

is relatively low in the bottom quantile of wildfire crime but increase sharply for the regions located in the top quantile. Therefore, the marginal effect on wildfire crime of one unit increase in inequality in the regions in the top quantile is much greater than the marginal effect in the regions that enjoy the more desirable position in the bottom quantile. Similarly, the marginal effect of an increase in unemployment and material deprivation is much higher in the top quantile, meaning that a unit increase in unemployment (or material deprivation) in regions with high unemployment (or material deprivation) has much greater impact on wildfire crime than a unit increase in the regions in the bottom quantile.

Looking at the effect of organised crime and homicide rates, once again the estimated coefficients for the top and bottom quantiles have positive signs and are significant. The estimated parameters for HOMR_{it} are particularly informative since the estimated coefficients for the top quantile are much greater in magnitude than those in the bottom quantile, thus indicating that the marginal effect of homicide rate on wildfire crime are much higher in regions with a high criminality level. This result highlights the fact that wildfire crime is strongly related to local criminal organisations, confirming the literature findings for other types of environmental crime (Germani et al. 2018).

Considering the effect of education on wildfire crime, it appears that EDUC_{it} has the expected negative signs throughout the estimated quantiles in all estimated models but model M1 (for the bottom quantile only), where the estimated parameters, is not significant. A comparison of the magnitude of the estimated coefficients in the bottom and top quantiles suggests that the marginal effects of the level of education increase in the top quantile, suggesting that an increase in one year in education in regions with higher wildfire crime has a greater impact than an extra year in education in regions with a lower level of wildfire crime. Looking at the estimated coefficients for UNIV_{it}, these are also significantly different from zero throughout the estimated models but the marginal effects are less clear.

Finally, focusing on the results for AGRI_{it}, the estimated coefficients are statistically significant for all estimated models. However, the estimated signs change from mostly negative in the bottom quantile to mostly positive in the top quantile. These results suggest that in the regions with a high level of employment in the agricultural sector, a further increase in employment also increases the rate of wildfire crime. On the contrary, in the regions in the bottom quantile, a rise in employment in agriculture is likely to provide a protective effect concerning wild forests. In the literature, the impact of this covariate is not clear and empirical works reports contradicting results. For example, Leone et al. (2002) found a positive relation between wildfire crime and agricultural expansion in the Mediterranean regions. The authors suggest that forests have been voluntarily set on fire to gain land for agriculture and pasture which were retained due to being more valuable than the land used for logging. Similarly, Moreira et al. (2009) describe burning shrubland as a widespread practice in rural areas across the Mediterranean region used to renovate livestock pastures. On the other hand, Martinez et al. (2009) suggest that land abandonment and rural exodus were found to be positively related to wildfire occurrence. To the best of our knowledge, the results of the quantile panel estimation bring a new perspective to the related literature as it suggests that the regions in bottom quantile may benefit from agricultural expansion. On the other hand, the regions in the top quantiles further expanding the possibility of employment in agriculture may induce greater wildfire occurrences. However, such a strong conclusion calls for further investigation. For example, the high degree of abandonment and depopulation of the high hill and mountain areas, and the abandonment of agronomic and silvicultural practices may have an influence on fire occurrence. The investigation of accidental wild forest fires is outside the scope of this work. However, the results in Table 3 warrant further investigation in future research.

4.1 Robustness checks

The reduced form estimations presented in Table 3 are based on the estimation of the quantile function $Q_{\mathrm{WF}_{it}}(p|R_{it})$. However, most of the related literature has considered the conditional distribution of $E\big(\mathrm{WF}_{i,t}|R_{i,t}\big)$. Also, the validity of the estimation results depends on the assumption of the validity of the instrument used. Accordingly, in this section, the estimation procedure in Table 3 is repeated but this time using the Poisson fixed effect panel model (for an excellent review, see Cameron and Trivedi 2013). Namely, we used the two-step GMM estimator with instrumental variables (IV-GMM). In the cross-section context, Poisson-type models have been used to investigate problems in criminology and criminal justice (see Osgood 2000 and the references therein). The advantage of using Poisson regression models is that the hypothesis of linearity is relaxed in the sense that a function of the mean of the observations is nonlinear in some of the sets of covariates. The hypothesis of normality is also relaxed to the assumption that the distribution belongs to the exponential family. In the context of longitudinal data, the application of Poisson-type models is still limited. However, Poisson fixed effect models have been used for predicting the number of human-caused wildfire in the work of Prestemon and Butry (2005), Levi and Bestelmeyer (2016), and Marchal et al. (2017).

For the IV-GMM Poisson procedure, under multiplicative errors, Eq. (1) becomes

$$\lambda_{it} = \exp\Big(\alpha_i + \beta_0' \tilde{R}_{it} + \beta_1' D_{it}\Big)\varepsilon_{it}. \tag{8}$$

where $\tilde{R}_{it}' = \Big\{\tilde{X}_{it}' Y_{it}'\Big\}$ are the exogenous regressors, D_{it} is the endogenous regressors, and ϵ_{it} is unit-mean errors. This leads to the following error function

$$u\Big(WF_{it}, \tilde{R}_{it}, D_{it}, \beta_0, \beta_1\Big) = \frac{WF_{it}}{\exp\Big(\alpha_i + \beta_0' \tilde{R}_{it} + \beta_1' D_{it}\Big) - 1}.$$

Let Π_{it}' be a vector of instrumental variables, such that the elements of Π_{it}' are correlated with the endogenous regressors in D_{it}. The population-moment condition for GMM estimation is

$$E\Big(\tilde{\Pi}_{it}, u(WF_{it}, \tilde{Z}_{it}, D_{it}, \beta_0, \beta_1\Big) = 0$$

where the vector Π'_{it} is partitioned as $\left(\tilde{Z}'_{it}, \Pi'_{it}\right)$. The sample-moment conditions are the sample analogues of the population-moment conditions, so that the GMM estimator solves the minimisation problem to make the sample-moment conditions as close to zero as possible (Cameron and Trivedi 2013; Wooldridge, 2010).

Table 4 presents the results for the IV-GMM estimation for the models considered in Table 3. Note that the estimation results were obtained using the same instruments used to produce the estimated parameters in Table 3. Upon comparing the estimation results in Table 3, the IV-GMM regressions suggest that socio-economic factors play an important role in explaining wildfire crime with most of the estimated coefficients being significantly different from zero. Overall, the estimated coefficients have the same signs but the IV-GMM estimated parameters are, generally speaking, smaller in magnitude than the estimated coefficients in Table 3. Therefore, there are no signs of misspecification in the estimated signs of the coefficients in Table 3.

Coming to misspecification tests, the consistency of the IV-GMM estimator depends on the validity instruments in the crime count regression. We address this issue by considering the Hansen J test of overidentifying restrictions, which tests the null hypothesis of the overall validity of the instruments by analysing the sample analogue of the moment conditions used in the estimation process. The test is asymptotically distributed as a χ^2-distribution with q degree of freedom. Failure to reject the null hypothesis gives support to the model under consideration. In addition to the Hansen J test, the assumption of serially uncorrelated structure of the error term was assessed by testing the null hypothesis that there is no serial correlation in the residuals using the Newey (1985) test. The test is based on the off-diagonal elements of the residual correlation matrix and is asymptotically χ^2-distributed. Failure to reject the null hypothesis of no first order autocorrelation indicates that the model is not well specified.

The bottom of Table 4 reports the p-values of the overidentification test. Looking at the results of the Hansen J test, in all cases, we fail to reject the null hypothesis that the model is correctly specified, supporting the validity of the instruments used in the model. Similarly, the Newey test fails to reject the null hypothesis of no first order autocorrelation, suggesting that the models are well specified.

4.2 Business cycles and wildfire crime

In his influential paper, Becker (1968) specified a crime production function where the decision to engage in crime is a negative function of the probability of being caught and a negative function of the income equivalent loss experienced by the offender for being caught and convicted. According to Becker's model, anything that raises the crime production cost lowers the expected utility of the crime. In this respect, we expect the opportunity cost of crime to be lower during the contraction phases of the business cycle. This is because GDP growth contractions are usually accompanied by an increase in unemployment, lower income, and greater deprivation. Also, reduced central and local income tax during the contraction periods may result in budget cuts for the crime-prevention policy which in turn reduces the probability of people being held accountable for crimes committed. Given the reduction in crime production costs,

Table 4 Estimation results for the two-step IV-GMM and CF-GMM models

Risk factors	M1 Coef	M2 Coef	M3 Coef	M4 Coef
$\ln INER_{it}$	0.101** (0.014)	0.010* (0.006)	0.110*** (0.014)	0.316*** (0.145)
$MATDEP_{it}$	0.249*** (0.072)	0.257*** (0.054)	–	–
$\ln UNEM_{it}$	–	–	0.517*** (0.237)	0.489* (0.286)
$\ln EDUC_{it}$	– 0.071 (0.017)	– 0.041*** (0.013)	–	–
$\ln UNIV_{it}$	–	–	– 0.053 (0.029)	– 0.249*** (0.045)
$\ln ORGC_{it}$	0.054** (0.030)	0.168*** (0.018)	0.296*** (0.104)	–
$\ln HOMR_{it}$	0.232*** (0.105)	0.320*** (0.080)	0.075* (0.041)	0.755** (0.030)-
$\ln INC_{it}$	- 0.961*** (0.091)	–	–	–
$\ln EMPL_{it}$	–	– 0.587*** (0.074)	–	–
$\ln TRIAL_{it}$	– 0.530*** (0.087)	– 0.421*** (0.080)	– 0.690*** (0.112)	– 0.731*** (0.1462)
$\ln EQI_{it}$	–	–	–	– 0.016*** (0.000)
$\ln DEN_{it}$	– 0.004*** (0.000)	– 0.002*** (0.000)	– 0.004*** (0.000)	– 0.001*** (0.001)
$\ln RAIN_{it}$	– 0.001** (0.000)	– 0.001*** (0.000)	– 0.001*** (0.000)	– 0.001*** (0.000)
$\ln TEMP_{it}$	0.047** (0.027)	0.061*** (0.022)	0.072*** (0.001)	0.144*** (0.037)
$\ln AGRI_{it}$	– 0.579*** (0.081)	0.518*** (0.156)	– 0.464*** (0.027)	– 0.078 (0.464)
CONST	11.406 (9.448)	10.633*** (1.3621)	0.861*** (0.068)	3.519*** (0.919)
Hansen	0.597	0.478	0.570	0.378
Serial corr	0.427	0.651	0.496	0.528

The table report the estimated coefficients of the Poisson two-step IV-GMM models (robust standard errors are presented in parentheses below their corresponding coefficients) with dependent variable WF_{it}
The Hansen J test is a test for overidentifying restriction. The instruments used are the covariates lagged and the interaction share inequality variable calculated as explained in Sect. 2.2. The "*ln*" indicates that the covariates were taken in natural logs
*, **, ***denote statistical significance at 10%; 5%; and 1%, respectively

we therefore expect a higher correlation between income and wildfire crime over the contraction periods and a lower correlation during the expansion periods.

Like other Southern European countries, Italy has undergone a prolonged recession period due to the sovereign debt crisis in 2009. The dip contraction in economic growth has led to an increase in inequality and poverty, hitting the bottom of the income distribution more severely. The southern regions were particularly affected by the economic crisis as they experienced a contraction in income transfer coupled with a sharp increase in unemployment rate. In this respect, a deterioration of economic fundamentals in the region may play a role in human-induced forest fires. A study by Leone et al. (2002) found that forests in the south of Italy were voluntarily set on fire to create firefighting jobs. Similarly, Lovreglio et al. (2010) reported that in the southern regions, fires were started by seasonal workers to force or maintain employment.

Against this background, a natural question arises: is wildfire crime related to the business cycle? In other words, do prolonged contractions in economic growth or sustained expansion periods have a relevant effect on wildfire crime patterns? The answers to these questions are important since the knowledge of crime behaviour patterns in relation to the economic cycle may inform environmental policymakers on the correct course of action to reduce the devastation caused by arson.

With this target in mind, we calculated the time-varying correlation coefficient between wildfire crime and income. The hypothesis we were testing was whether a contraction in income (and other economic fundamentals such as unemployment) reduced the crime production cost. As Becker's theoretical model suggests, we should see the correlation between income and wildfire crime escalate during contraction periods of the business cycle. Also, if a lower opportunity cost increases the probability of crime, we should expect a greater persistence in the correlation over time in regions with lower income (greater unemployment level).

To investigate this hypothesis, we first ranked each region in the sample according to the degree of wildfire crime. To do so, we considered the quantile distribution function of the wildfire distribution for each region over time. Namely, each region i at time t was classified as belonging to the 25th, 50th, 75th, or 100th percentile of wildfire distribution by assigning a code from 1 to 4 to each percentile, with the 25th percentile being code 1 and the 100th percentile being code 4. In doing so, we were able to classify each region in each year as belonging to a given group according to the code assigned. Finally, the twenty Italian regions were classified into four groups according to the gravity of the wildfire crime occurrence. Table 8 in the Appendix reports the classification of the regions by code. It shows that with the noticeable exception of the northern region of Liguria, all regions in the top quantile were in the south of Italy where the per capita income is the lowest. Figure 1 reports, for each group of regions, the average per capita income, inequality, and material deprivation over time. From Fig. 1, it appears that the ranking of the Italian regions in term of wildfire crime exactly matches the level of per capita income.

Figure 2 shows the time-varying correlation coefficients between $\mathrm{WF_t}$ and $\mathrm{INC_t}$ by code. The correlation coefficients for $\mathrm{INEQ_t}$ and $\mathrm{MATDEP_t}$ were also calculated for completeness. In particular, Fig. 2a–c report, by region code, the time-varying correlation coefficients between $\mathrm{WF_t}$ and $\mathrm{INC_t}$, $\mathrm{INEQ_t}$,and $\mathrm{MATDEP_t}$, respectively.

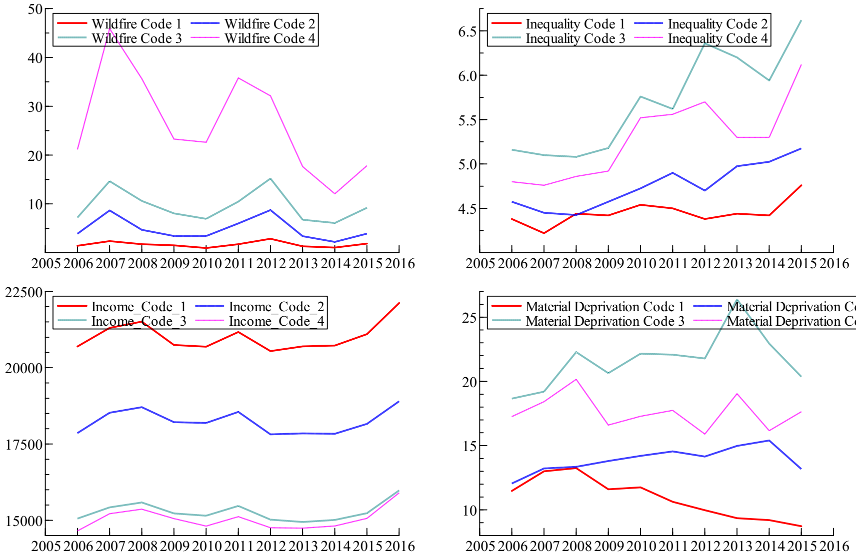

Fig. 1 Plots of wildfire, per capita income, material deprivation, and inequality during the sample under consideration by region code points. See Table 8 for code classification

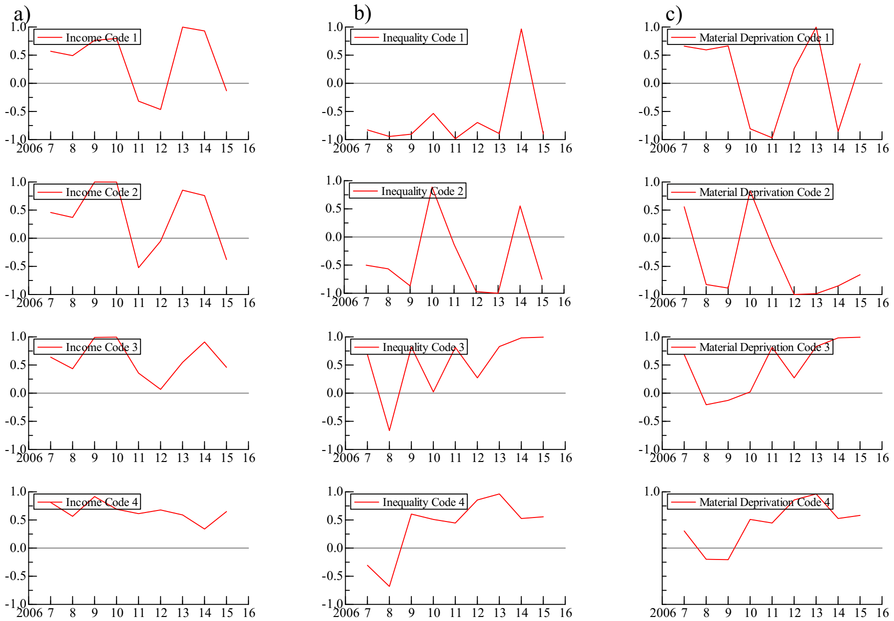

Fig. 2 Time-varying correlation coefficients by region code between wildfire and per capita income, inequality, material deprivation, respectively. *Note* Time-varying correlation coefficients in the *y*-axes. See Table 8 for code classification

For each code, the time-varying correlation coefficients were calculated using a 3 year-moving window as follows:

$$\rho_{\mathrm{t}} = \frac{\sum_{\mathrm{t}-8}^{T}\left(\mathrm{WF_t} - \overline{\mathrm{WF_t}}\right)\left(\mathrm{INC_t} - \overline{\mathrm{INC_t}}\right)}{\left(\sqrt{\sum_{\mathrm{t}-8}^{T}\left(\mathrm{WF_t} - \overline{\mathrm{WF_t}}\right)^2}\right)\left(\sqrt{\sum_{\mathrm{t}-8}^{T}\left(\mathrm{WF_t} - \overline{\mathrm{WF_t}}\right)^2}\right)}.$$

Figure 2a shows that the correlation between $\mathrm{WF_t}$ and $\mathrm{INC_t}$ changes over time quite substantially, no matter the region code taken into consideration. There is clear evidence that the recession that started in 2009 had an impact on the wildfire crime rate in Italy. Indeed, regardless of the region code under consideration, the correlation coefficients in Fig. 2a are positive and greater in magnitude in terms of the correspondence of income through 2010 to 2014, confirming the predictions of the theoretical model proposed by Becker (1968). Also, it appears that the opportunity cost mechanism has a greater impact on the regions classified in Codes 3 and 4, which have a lower income. For these regions, the correlation coefficient is mostly positive. Also in these regions, the calculated coefficients ρ_{t} show a greater persistence over time and fail to revert the negative sign over the period under consideration. On the other hand, it appears that for the regions in Codes 1 and 2, the signs of ρ_{t} show the features of a mean reverting process. The calculated coefficients in Fig. 2b, c have similar patterns.

To summarise the results in this section, a simple analysis of the time-varying correlation coefficients reveals that economic downturns have a significant impact on the probability of the type of environmental crime rate considered in this paper. These results are in line with the literature where it was found that the business cycle has a range of important effects on the crime rate (Grogger 1998; Machin and Meghir 2004).

5 Conclusion and policy implications

The empirical investigation presented in this work complements the literature on environmental crime by providing new evidence on the determinants of human-induced wildfire. In particular, by using panel data at the regional level for the period between 2006 and 2015, this study provides new insights into the role of socio-economic factors when explaining wildfire crime in Italy.

The contribution of this paper to the extant literature can be summarised as follows. First, it conducts an extensive analysis of the relationship between socio-economic determinants and wildfire crime. In the literature, empirical works on the socio-economic factors influencing wildland arson are still quite limited. However, understanding the socio-economic drivers behind wildfires are crucial for the planning and management of fire prevention policies. In this respect, the empirical investigation carried out in this work reveals that socio-economic factors are important determinants of wildfire crime in Italy. In particular, the estimation results presented in this paper clearly suggest that higher levels of material deprivation, unemployment and income inequality increase wildfire crime. Also, risk factors such as organised crime and homicide were seen to affect the probability of crime in the same direction. On the other hand, a negative relationship between the level of education and wildfire crime was found. These findings are important since, according to the Italian environmental group Legambiente (2010), more than half of Italy's fires are started deliberately.

The proposed methodological approach is the second contribution of this paper. Most of the related literature uses linear or nonlinear regression models to investigate the relationship between human-induced wildfire crime and the risk factors by modelling the conditional mean function of the wildfire distribution (see for example, Mancini et al. 2018; Pazienza and Beraldo 2004; Michetti and Pinar 2019). However, the conditional mean modelling framework used in these works has some important limitations since it cannot be readily extended to the noncentral locations of the wildfire distribution. In this paper, we are interested in understanding the heterogeneity of the relationship across the wildfire distribution, in addition to the relationship at the mean or median. Accordingly, the adopted quantile panel model allows the investigator to move away from the central moment and consider the tails of the distribution where Italian regions with the highest and least levels of crime are located. This different perspective is certainly interesting from a policy point of view, since policymakers are more interested in knowing what happens at the extremes of a distribution than at the centre. The estimation results suggest that the impact of a one unit increase in inequality in the regions in the top quantile (where the inequality level is the highest) is much greater than the marginal effect in the regions that enjoy the desirable position at the bottom quantile, (where the level of inequality is the lowest). Similarly, the marginal effects of increases in unemployment and material deprivation are much higher for the regions in the top quantile, meaning that a unit increase in unemployment (or material deprivation) in these regions has a much greater impact on wildfire crime than a unit increase in the regions in the bottom quantile.

The findings from the quantile panel and Poisson models also highlight the fact that wildfire crime data cannot be easily estimated using models that do not consider fat tails, such as ordinary least-squared regressions. This is because the least-squared methods do not have a bounded influence and cannot accommodate for heterogeneity, which is an important characteristic of the Italian data. Heavy tails and outliers can make the estimators inconsistent and, even when the estimator itself is consistent, the standard errors require higher moments to exist (Cameron and Trivedi 2013). In this context, the focus on the conditional mean may become a misleading measure of the central location because it is heavily influenced by outliers. In this respect, the quantile regression method adopted in this paper is robust in terms of departures from the normality of innovations assumption (Chernozhukov and Hansen 2008).

Third, this study takes a novel approach with respect to the extant literature by investigating the relationship between the business cycle and wildfire crime. The well-established literature on business cycles advocates the viewpoint that the type of socio-economic risk factors considered in the paper is all overwhelmingly cyclical. We therefore speculate that wildfire crime patterns over time may also stem from the properties of socio-economic determinants. It was found that per capita income and wildfire crime occurrence share similar patterns over time. In addition, the analysis of the time-varying correlation coefficients reveals that regions in the top quantile distribution of the wildfire distribution (regions with higher level of wildfire crime) have also mostly positive estimated time-varying correlation coefficients between socio-economic risk factors and wildfire.

The joint interpretation of the empirical results presented in this paper provides several insights that may be useful when designing more effective policy interventions designed to reduce the risks of forest fire crime and their environmental impact. First, the overall findings in this paper confirm the strong correlation between risk factors such as inequality, poverty and unemployment and wildfire crime as was found in the related literature. These results emphasise the need for a strong policy framework to reduce income inequality and promote economic growth which in turn reduces unemployment and poverty. Consensus in the general crime literature suggests that a high crime rate hampers standard of living and overall quality of life (Coccia 2017; Grogger and Michael 2000). This literature also suggests that there is a potential vicious cycle between crime unemployment and poverty. Prevalent criminal activities erode employment opportunities and are exacerbated by the high unemployment rates. In this respect, social policies should be formulated to strengthen the equal distribution of income which will ultimately reduce the wildfire crime rates and poverty incidence across the country, especially in the Italian southern regions.

Second, the quantile panel estimation results suggest that level of education can affect wildfire crime. Regions characterised by a lower education level show a higher wildfire crime risk. Although the relation between wildfire risk and education is still an open question in the wildfire literature (see, for example, Lochner 2004; Lochner and Moretti 2004; Hjalmarsson 2008; Oreopoulos 2006; Lochner 2004; Usher 1997), it is largely accepted that level of education plays an important role in determining the distribution of income: a higher level of educational attainment improves the income level and reduces income inequality and poverty (see, for example, Glaeser et al. 2004). In this respect, a higher education level may affect wildfires indirectly by reducing the income inequality and increasing the income, resulting in less incentive to engage in environmental crime.

Third, the estimated negative signs for the trial rate suggest that there is a negative relationship between wildfire crime occurrence and level of apprehension. An inspection of the magnitude of the estimated coefficients across the quantiles reveals that regions with a higher level of law enforcement also experience lower levels of wildfire crime occurrence. Italy is also characterised by the presence of organised crime which controls the economic activities connected to agriculture and pasture practices. The findings in this paper corroborate the view that the presence of illegal associations increases the occurrence of wildfire crime. Taken together, these results suggest that human-induced wildfire occurrences may benefit from a greater level of law enforcement. In this respect, the wildfire crime deterrence law in Italy establishes temporal binding constraints on the use of the area damaged by forest fire. However, evidence suggests that these regulations have only been partially enforced in some Italian regions (the EFFACE 2016 report). In this respect, a greater enforcement of the deterrence regulations may reduce the wildfire crime occurrence. The issue of increasing law enforcement and other forms of crime apprehension has become particularly urgent after the Covid-19 pandemic, since Italy has received a large post-pandemic economic recovery fund from the European Union as support towards the green transition which may attract criminal organisations looking to benefit from lucrative reforestation contracts to repurpose the land for solar panels.

Overall, the results presented in this study enhance our understanding of the relationship between socio-economic risk factors and wildfires. A possible limitation of this study is that the use of data at an annual frequency may conceal the impact of weather-related causes of wildfire, which is overwhelmingly seasonal. Weather elements are undeniably recognised as major determinants of wildfire risk in the relevant literature (see Ganteaume et al. 2013; and the references therein). Previous studies suggest that socio-economic factors are determinants of the fire frequency, whereas the surface burned per fire is more related to weather patterns (Michetti and Pinar 2019). In this respect, a robustness check may be undertaken in the future to check whether increasing the frequency of the data still confirms the finding of this paper. Another possible limitation of this study is that the analysis carried out only relates to wildfire crime in Italy. However, there are other countries in Europe that are plagued by a high rate of wildfire crime and have a great socio-economic divide such as Spain, for example (see Fig. 3 in the Appendix). Therefore, our future research will explore whether the risk factors that we have identified for Italy also hold for other countries in a similar situation.

Funding Open access funding provided by Università degli Studi di Torino within the CRUI-CARE Agreement.

Declarations

Conflict of interest The author has no relevant financial or non-financial interests to disclose. The author has no competing interests to declare that are relevant to the content of this article. The author certifies that they have no affiliations with or involvement in any organisation or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The author has no financial or proprietary interests in any material discussed in this article.

Appendix

See Fig. 3 and Tables 5, 6, 7, and 8.

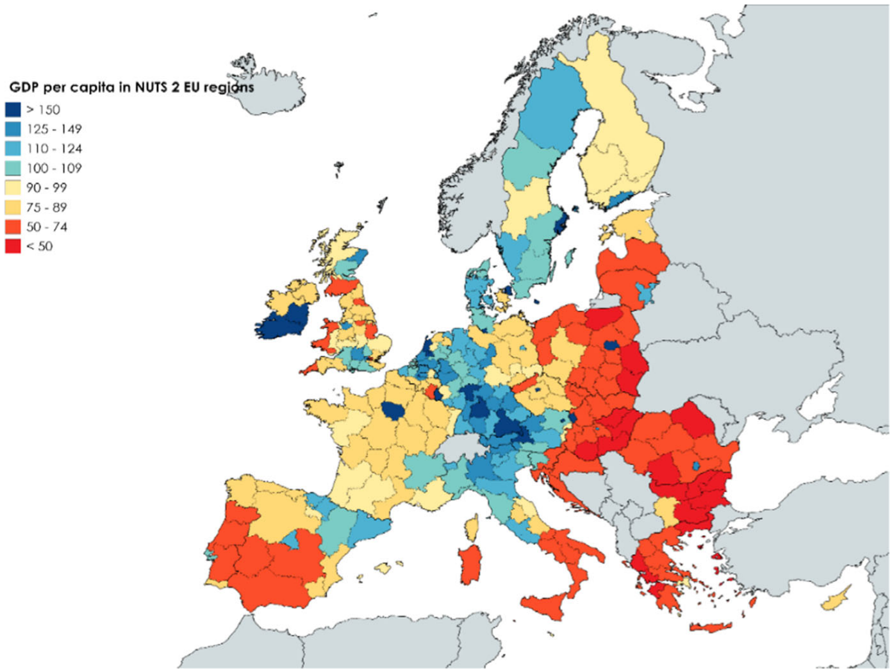

Fig. 3 Gross domestic product (GDP) per inhabitant in purchasing power standards in relation to the EU-28 average by European region (NUTS 2) in 2017. *Source* Eurostat

Table 5 Correlation matrix

	WF_{it}	$INER_{it}$	$MATDEP_{it}$	$UNEM_{it}$	$EDUC_{it}$	$UNIV_{it}$	$ORGC_{it}$	$HOMR_{it}$	INC_{it}	$EMPL_{it}$	$TRIAL_{it}$	DEN_{it}	$RAIN_{it}$	$TEMP_{it}$	$AGRI_{it}$
WF_{it}	1														
$INER_{it}$	0.29***	1													
$MATDEP_{it}$	0.53***	0.57***	1												
$UNEM_{it}$	0.36***	0.71***	0.64***	1											
$EDUC_{it}$	− 0.32***	− 0.47***	− 0.33***	− 0.57***	1										
$UNIV_{it}$	− 0.30***	− 0.18**	− 0.18**	− 0.29***	0.74***	1									
$ORGC_{it}$	0.35***	0.22*	0.30***	0.25***	− 0.17**	− 0.06	1								
$HOMR_{it}$	0.60***	0.41***	0.46***	0.43***	− 0.48***	− 0.30***	0.283***	1							
INC_{it}	− 0.57***	− 0.55***	− 0.65***	− 0.80***	0.68***	0.57***	− 0.375***	− 0.48***	1						
$EMPL_{it}$	− 0.42***	− 0.42***	− 0.31***	− 0.32***	0.13*	0.16**	− 0.289***	− 0.15*	0.54***	1					
$TRIAL_{it}$	0.05	0.06	0.05	0.15**	− 0.04	0.03	0.074	0.09	− 0.18**	0.04	1				
DEN_{it}	− 0.33***	0.26**	− 0.08	− 0.08	0.03	0.18**	− 0.076	0.04	0.31***	0.73***	0.13	1			
$RAIN_{it}$	− 0.31***	− 0.24**	− 0.26***	− 0.32***	0.46***	0.32***	− 0.130*	− 0.23***	0.35***	0.11	− 0.05	0.20**	1		
$TEMP_{it}$	0.44***	0.52***	0.53***	0.65***	− 0.52***	− 0.27***	0.265***	0.39***	− 0.72***	− 0.41***	0.27***	− 0.18**	− 0.41***	1	
$AGRI_{it}$	0.65***	0.39***	0.53***	0.68***	− 0.63***	− 0.58***	0.248***	0.53***	− 0.84***	− 0.46***	0.13*	− 0.50***	− 0.40***	0.60***	1

*, **, ***denote statistical significance at 10%; 5%; and 1%, respectively

Table 6 Principal component analysis of two variables: "Proportion of household in economic distress" and "Proportion of household in material deprivation"

Variable	Proportion
Component 1	0.8876
Component 2	0.1124

The table contains the principal component analysis (PCA) of the two covariates. The PCA analysis involves a linear transformation of the data in such a way that all principal components combined contain the same information as the original variables. However, the important information is partitioned over the components in such way that the components are orthogonal, and earlier components contain more information than later components. In Table 6, the first principal component has maximal overall variance and explain approximately 89% of the total variance. The second principal component has variance approximately 11%

Table 7 Condition index for the estimated models in Table 3

	M1	M2	M3	M4
$\ln\text{INER}_{it}$	2.37	2.36	2.68	4.98
MATDEP_{it}	2.41	2.03	–	–
$\ln\text{UNEM}_{it}$	–	–	3.78	– 4.82
$\ln\text{EDUC}_{it}$	3.34	2.86	–	–
$\ln\text{UNIV}_{it}$	–	–	1.90	2.97
$\ln\text{ORGC}_{it}$	1.31	1.28	1.14	2.24
$\ln\text{HOMR}_{it}$	1.68	1.65	1.59	–
$\ln\text{INC}_{it}$	8.12	–	–	–
$\ln\text{EMPL}_{it}$	–	2.92	–	–
$\ln\text{TRIAL}_{it}$	1.23	1.21	1.11	1.35
$\ln\text{EQI}_{it}$	–	–	–	1.15
$\ln\text{DEN}_{it}$	2.93	4.50	2.83	3.98
$\ln\text{RAIN}_{it}$	1.47	1.42	1.30	1.66
$\ln\text{TEMP}_{it}$	2.49	2.22	2.12	4.30
$\ln\text{AGRI}_{it}$	6.52	4.77	5.42	5.18

Table 8 Classification of Italian regions by quantile according to the degree wildfire crime rate

Code 1	Code 2	Code 3	Code 4
Trentino Alto Adige	Valle Aosta	Umbria	Molise
Veneto	Marche	Toscana	Basilicata
Emilia Romagna	Friuli Venezia Giulia	Puglia	Liguria
Piemonte	Abruzzo	Sicilia	Sardegna
Lombardia	Lazio	Campania	Calabria

Codes 1, 2,3 and 4 correspond to the 25th, 50th, 75th and 100th percentile of wildfire distribution (WF), respectively

References

A'Hearn B, Venables AJ (2013) Regional disparities: internal geography and external trade. In: Toniolo G (ed) The Oxford handbook of the Italian economy since unification. Oxford University Press, Oxford, pp 599–630

Albertson K, Aylen J, Cavan G, McMorrow J (2009) Forecasting the outbreak of Moorland wildfires in the English Peak district. J Environ Manag 90:2642–2651

Aldieri L, Vinci C (2017) Quantile regression for panel data: an empirical approach for knowledge spillovers endogeneity. Int J Econ Financ 9:16–113

Bartik TJ (1991) Who benefits from state and local economic development policies? W. E. Upjohn Institute for Employment Research

Becker G (1968) Crime and punishment: an economic approach. J Polit Econ 76:169–217

Blanchard OJ, Katz LF, Hall RE, Eichengreen B (1992) Regional evolutions. Brook Pap Econ Act 1:1–75

Bourguignon F (2000) Crime, violence, and inequitable development. In: Pleskovic B, Stiglitz JE (eds) Annual World Bank conference on development economics 1999. Washington, D.C.

Boustan L, Ferreira F, Winkler H, Zolt E (2013) The effect of rising income inequality on taxation and public expenditures: evidence from US municipalities and school districts, 1970–2000. Rev Econ Stat 95:1291–1302

Bushway S, Phillips M, Cook PJ (2019) The overall effect of the business cycle on crime. German Econ Rev 13:436–446

Cameron AC, Trivedi PK (2013) Regression analysis of count data, 2nd edn. Cambridge University Press, New York

Canepa A, Drogo F (2021) Wildfire, apprehension and social vulnerability in Italy. For Policy Econ 122:103230

Carvalho A, Flannigan MD, Logan K, Miranda AI, Borrego C (2008) Fire activity in Portugal and its relationship to weather and the Canadian fire weather index system. Int J Wildland Fire 17:328–338

Catry FX, Damasceno P, Silva JS, Galante M, Moreira F (2007) Spatial distribution patterns of wildfire ignitions in Portugal. In: Conference wildfire 2007, Seville, Spain

Charron N, Lapuente V, Annoni P (2019) Measuring quality of government in EU regions across space and time. Pap Reg Sci 98:1925–2195

Chernozhukov V, Hansen C (2008) Instrumental variable quantile regression: a robust inference approach. J Econom 142:379–398

Coccia M (2017) A theory of general causes of violent crime: homicides, income inequality and deficiencies of the heat hypothesis and of the model of CLASH. Aggress Violent Behav 37:190–200

Cook PJ, Durrance CP (2013) The virtuous tax: lifesaving and crime-prevention effects of the 1991 federal alcohol-tax increase. J Health Econ 32:261–267

Cook PJ, Moore MJ (1993) Drinking and schooling. J Health Econ 12:411–429

Di Fonzo M, Falcone PM, Germani AR, Imbriani C, Morone P, Reganati F, (2015) Impacts of forest fire crimes. Report compiled as part of the EFFACE project, University of Rome

Eck J, Maguire E (2000) Have changes in policing reduced violent crime? An assessment of the evidence. In: Blumstein A, Wallman J (eds) The crime drop in America. Cambridge University Press, New York

EFFACE (2016) Final report summary. European Union Action to Fight Environmental Crime

Ehrlich I (1973) Participation in Illegitimate activities: a theoretical and empirical investigation. J Polit Econ 81:521–565

Enamorado T, López-Calva LF, Rodríguez-Castelán C, Winkler H (2016) Income inequality and violent crime: evidence from Mexico's drug war. J Dev Econ 120:128–143

Fajnzylber P, Lederman D, Loayza N (2002) Inequality and violent crime. J Law Econ 45:1–40

Ganteaume A, Camia A, Jappiot M, San-Miguel-Ayanz J, Long-Fournel M, Lampin C (2013) A review of the main driving factors of forest fire ignition over Europe. Environ Manag 51:651–662

Germani AR, Pergolizzi A, Reganati F (2018) Eco-mafia and environmental crime in Italy: evidence from the organised trafficking of waste. In: Spapens T, White R, van Uhm D, Huisman W (eds) Green crimes and dirty money. Taylor and Francis Routledge Group, London

Glaeser E, La Porta R, Shleifer A (2004) Do institutions cause growth? J Econ Growth 9:271–303

Gonzalez-Olabarria JR, Mola-Yudego B, Coll L (2015) Different factors for different causes: analysis of the spatial aggregations of fire ignitions in Catalonia (Spain). Risk Anal 35:1197–1209

Gould ED, Weinberg BA, Mustard DB (2002) Crime rates and local labor market opportunities in the United States: 1979–1997. Rev Econ Stat 84:45–61

Grogger J (1998) Market wages and youth crime. J Law Econ 16:756–791
Grogger J, Michael W (2000) The emergence of crack cocaine and the rise in urban crime rates. Rev Econ Stat 82:519–529
Grossman HI (1991) A general equilibrium model of insurrections. Am Econ Rev 81:912–921
Guo F, Wang G, Su Z, Liang H, Wang W, Lin F, Liu A (2016) What drives forest fire in Fujian, China? Evidence from logistic regression and random forests. Int J Wildland Fire 25:505–519
Harris G, Vermaak C (2015) Economic inequality as a source of interpersonal violence: evidence from Sub-Saharan Africa and South Africa. S Afr J Econ Manage Sci 18:45–87
Hjalmarsson R (2008) Criminal justice involvement and high school completion. J Urban Econ 63:613–630
Imrohoroglu A, Merlo A, Rupert P (2000) On the political economy of income redistribution and crime. Int Econ Rev 41:1–25
Johansson E, Bockerman P, Prattala R, Uutela A (2006) Alcohol-related mortality, drinking behavior, and business cycles: are slumps really dry seasons? Eur J Health Econ 7:215–220
Koenker R (2004) Quantile regression for longitudinal data. J Multivar Anal 91:74–89
Legambiente, Rapporto Ecomafie 2010, Rome
Leone V, Lovreglio R, Martinez-Fernandez J (2002) Forest fires and anthropic influences: a study case (Gargano National Park, Italy). In: Viegas X (ed) Forest fire research and wildland fire safety. Mill Press, Rotterdam, pp 11–28
Levi MR, Bestelmeyer BT (2016) Biophysical influences on the spatial distribution of fire in the desert grassland region of the Southwestern USA. Landscape Ecol 31:2079–2095
Lochner L (2004) Education, work and crime: a human capital approach. Int Econ Rev 45:811–843
Lochner L, Moretti E (2004) The effect of education on crime: evidence from prison inmates, arrests, and self-reports. Am Econ Rev 94:155–189
Lovreglio R, Leone V, Giaquinto P, Notarnicola A (2010) Wildfire cause analysis: four case-studies in southern Italy. iForest 3:8–15
Machin S, Meghir C (2004) Crime and economic incentives. J Hum Resour 39:958–979
Maingi JK, Henry MC (2007) Factor influencing wildfire occurrence and distribution in Eastern Kentucky, USA. Int J Wildland Fire 16:23–33
Mancini LD, Corona P, Salvati L (2018) Ranking the importance of wildfires' human drivers through a multi-model regression approach. Environ Impact Assess Rev 72:77–186
Marchal J, Cumming SG, McIntire EJB (2017) Exploiting poisson additivity to predict fire frequency from maps of fire weather and land cover in boreal forests of Quebec, Canada. Ecography 40:200–209
Martınez J, Vega-Garcıa C, Chuvieco E (2009) Human-caused wildfire risk rating for prevention planning in Spain. J Environ Manag 90:1241–1252
Mehlum H, Miguel E, Torvik R (2006) Poverty and crime in 19th century Germany. J Urban Econ 59:370–388
Mercer DE, Prestemon JP (2005) Comparing production function models for wildfire risk analysis in the wildland urban interface. For Policy Econ 7:782–795
Michetti M, Pinar M (2019) Forest fires across Italian regions and implications for climate change: a panel data analysis. Environ Resource Econ 72:207–246
Moreira F, Vaz P, Catry F, Silva JS (2009) Regional variations in wildfire susceptibility of land-cover types in Portugal: implications for landscape management to minimize fire hazard. Int J Wildland Fire 18:563–574
Musolino D (2018) The north–south divide in Italy: reality or perception? Eur Spat Res Policy 25:29–53
Neumayer E (2005) Inequality and violent crime: evidence from data on robbery and violent theft. J Peace Res 42:101–112
Newey WK (1985) Maximum likelihood specification testing and conditional moment tests. Econometrica 53:1047–1070
Oreopoulos P (2006) Estimating average and local average treatment effects of education when compulsory schooling laws really matter. Am Econ Rev 96:152–175
Osgood WD (2000) Poisson-based regression analysis of aggregate crime rates. J Quant Criminol 16:21–43
Padilla M, Vega-Garcia C (2011) On the comparative importance of fire danger rating indices and their integration with spatial and temporal variables for predicting daily human-caused fire occurrences in Spain. Int J Wildland Fire 20:46–58
Pazienza P, Beraldo S (2004) Adverse effects and responsibility of environmental policy: the case of forest fires. Corp Soc Responsib Environ Manag 11:222–231

Pena J, Renegar J (2000) Computing approximate solutions for convex conic systems of constraints. Math Program 87:351–383

Plucinski MP (2014) The timing of vegetation fire occurrence in a human landscape. Fire Saf J 67:42–52

Powell D (2020) Quantile treatment effects in the presence of covariates. Rev Econ Stat 102:1–39

Powell D (2022) Quantile regression with nonadditive fixed effects. Empir Econ 63:2675–2691

Preisler HK, Brillinger DR, Burgan RE, Benoit JW (2004) Probability based models for estimation of wildfire risk. Int J Wildland Fire 13:133–142

Prestemon JP, Butry DT (2005) Time to burn: modeling wildland arson as an autoregressive crime function. Am J Agric Econ 87:756–770

Ricotta C, Guglietta D, Migliozzi A (2012) No evidence of increased fire risk due to agricultural land abandonment in Sardinia (Italy). Nat Hazard 12:1333–1336

Rodrigues M, Jimenez A, de la Riva J (2016) Analysis of recent spatial-temporal evolution of human driving factors of wildfires in Spain. Nat Hazards 84:2049–2070

Romero-Calcerrada R, Novillo CJ, Millington JDA, Gomez-Jimenez I (2008) GIS analysis of spatial patterns of human-caused wildfire ignition risk in the SW of Madrid (Central Spain). Landscape Ecol 23:341–354

Savona EU, Riccardi M (2018) Mapping the risk of serious and organised crime infiltration in european businesses—final report of the MORE Project. Milano: Transcrime, Università Cattolica del Sacro Cuore

Sebastian-Lopez A, Salvador-Civil R, Gonzalo-Jimenez J, San-Miguel-Ayanz J (2008) Integration of socio-economic and environmental variables for modelling long-term fire danger in Southern Europe. Eur J For Res 127:149–163

Seter H (2016) Connecting climate variability and conflict: implications for empirical testing. Polit Geogr 53:1–9

Tauchen H, Witte A, Griesinger H (1994) Criminal deterrence: revisiting the issue with a birth cohort. Rev Econ Stat 76:399–412

Torres De, Curth M, Biscayart C, Ghermandi L, Pfister G (2012) Wildland-urban interface fires and socioeconomic conditions: a case study of a Northwestern Patagonia city. Environ Manag 49:876–891

Usher D (1997) Education as a deterrent to crime. Can J Econ 30:367–384

Vasilakos C, Kalabokidis K, Hatzopoulos J, Matsinos I (2009) Identifying wildland fire ignition factors through sensitivity analysis of a neural network. Nat Hazards 50:125–143

Vélez R (2009) The causing factors: a focus on economic and social driving forces. In: Living with wildfires: what science can tell us. A contribution to the science. European Discussion Papers. EFI Discussion Paper 15