

Homicide Clearance: Discretionary and Non-discretionary Factors in Italy

Abstract

Previous studies have produced mixed findings regarding discretionary and non-discretionary factors associated with the likelihood of homicides being cleared. Performing Pearson's χ^2 test, logistic regressions, and random forest analyses on all homicide cases that occurred in Italy in 2014, we assess if factors pertaining to the discretionary domain –i.e., nationality, age, sex, socioeconomic status of the victim, previous convictions– and non-discretionary factors –e.g., weapon, location, circumstance– are correlated with the outcome of the investigation. The empirical analysis does not identify relations between victim's nationality, socioeconomic status, previous criminal records and homicide clearance. On the other hand, homicides whose victim is male are less likely to be cleared. While high media coverage correlates with a higher clearance rate, low media coverage is not correlated with the homicide case remaining unsolved. Finally, especially in Southern Italy where mafia organizations are more entrenched, homicides committed in a criminal environment have a lower chance of being cleared. With respect to Italy, the results indicate a lack of support for the discretionary perspective, that argues that police may use discretion in clearing homicide cases according to the sociodemographic characteristics of victims.

Keywords

Clearance Rate; Homicide Investigation; Murder Case; Theory of Law; Random Forest

Introduction

The homicide investigation resolution rate is a measure of the overall effectiveness of a police force. Failure to identify and apprehend those who commit homicides has a negative impact on society since it reduces the deterrent effect of the criminal justice system and public trust in the police (Nagin, 1998; Riedel & Jarvis, 1999). Moreover, homicides that remain unsolved may trigger new homicides as acts of revenge, especially if they take place in a criminal setting (Liem et al., 2019). For these reasons, it is important to gain insight into what factors determine homicide clearance and how discretionary action by criminal investigators may impact on the clearance of specific murder cases. In turn, deeper understanding of the relationship between discretion and clearance would support decision-makers in enhancing the effectiveness of homicide investigators and making societies more equal and fairer. At the same time, identifying what factors relate to the capacity of police forces to solve homicides would help them to allocate resources in a more efficient and effective manner.

Understanding of the factors, which influence the resolution of homicide cases is still limited in several respects. First, much of the literature on homicide clearance has focused on the United States – more generally, on the English-speaking countries– and provides partially contrasting results with respect to the relevance of multiple factors (e.g., Addington, 2007; Alderden & Lavery, 2007; Braga & Dusseault, 2018; Campedelli, 2022; Hawk & Dabney, 2014; Hough et al., 2019; Jiao, 2007; Lee, 2005; Litwin, 2004; Mancik et al., 2018; Ousey & Lee, 2010; Pizarro et al., 2020; Puckett & Lundman, 2003; Roberts, 2007; Roberts & Lyons, 2009, 2011). Second, few studies have efficaciously considered and operationalized the interaction of victims' characteristics, such as victim's socio-economic condition, which might trigger discretionary actions by police forces, and non-discretionary factors, such as weapon used to commit the homicide (Petersen, 2017; Vaughn, 2020).

In this study, we investigate factors associated with the clearance of Italian homicide cases and their interactions. Analysis of the Italian context is important for three reasons. Given the decline in homicides in most societies and countries in the world, as observed by LaFree et al. (2015), Italy, which has a low homicide rate –about 0.7 homicides per 100,000 inhabitants– provides a glimpse of what may occur in other countries in the future. At the same time, Italy, like other Western societies, is experiencing a migration inflow which exacerbates ethnic rivalries and fosters xenophobic sentiments (Amin, 2013). Finally, in Italy they operate structured organized crime groups which have a code of silence on criminal activity and a refusal to give evidence to the police –*omertà*– among their key rules of conduct (Paoli, 2003). In this regard, study of the Italian case furnishes insights into the possible correlates between a discretionary approach to murder investigation, homicide characteristics, and the presence of governance-type organized crime.

In view of the above, this study contributes to the broader literature on homicide clearance by: i) compensating for the relative dearth of studies of this type outside the United States; ii) conducting a country-based multisite study made possible by the fact that Italy has police forces which operate across the entire country; iii) analyzing both big metropolitan and small provincial investigative units, while previous research mainly concerned the former, which had managed multiple murder cases per year; iv) distinguishing between easy-to-solve and difficult-to-solve cases on the basis of the nature of the investigative unit handling the case; v) considering interaction of discretionary and non-discretionary factors.

The paper is structured as follows. The next section provides an overview of the theories and literature on homicide clearance. In the following section, we outline the purpose of the current study; then, we present the empirical strategy adopted and the data exploited. The results section shows that the outcome of the investigation is primarily related to the complexity of the case. The gender of the victim

is the only factor pertaining to the discretionary domain to have consistently a relationship with clearance rate. These results are then discussed considering the previous literature on the homicide clearance and Italian specificities.

Perspectives on homicide clearance

Within the field of studies on homicides and clearances, scholars propose two main contrasting interpretations on discretionary action by investigators: ‘discretionary’ vs. ‘non-discretionary’ perspectives. These theoretical approaches are then integrated by police devaluation and victim lifestyle theories. Scholars referring to the discretionary perspective, also known as ‘victim devaluation’ or ‘extralegal’ perspective, suggest that police forces modulate their efforts to solve homicide cases based on sociodemographic characteristics of the victims (Black, 1976; Paternoster, 1984; Petersen, 2017; Peterson & Hagan, 1984). According to this ‘discretionary’ perspective, factors such as the social position or demographic characteristics of the victim affect the interaction between citizens and the judicial and executive powers, and, in turn, impact upon the likelihood of a homicide case being cleared. In particular, the sex and gender, age, race, and socioeconomic status of the victim can affect clearance rates.

Related to the discretionary perspective is the ‘police devaluation’ perspective, which assumes a more sociological viewpoint and focuses on mistrust in the police by citizens with socioeconomically disadvantaged backgrounds. The negative perception of the police hinders the development of a positive relationship between disadvantaged, minority individuals and neighborhoods and law enforcement agencies. In turn, this attitude makes the investigative work harder and consequently negatively affects clearance in disadvantaged neighborhoods because it reduces the engagement of potential witnesses, who choose not to cooperate with the homicide investigation (Keel et al., 2009; Litwin, 2004; Litwin & Xu,

2007; Pizarro et al., 2020; Riedel & Jarvis, 1999). On this basis, Rydberg and Pizarro (2014) developed the so-called ‘victim lifestyle’ perspective, which maintains that victims’ past behaviors, particularly their involvement in criminal activities, can hinder clearance. This is due to (a) the reduced opportunities to rely on cooperative witnesses –either because of fear of retaliation or so as not to be labelled a ‘snitch’– and (b) less cooperation by victims’ relatives –because they are committed to preserving their own criminal activities, even those not related to the murder itself.

By contrast, the non-discretionary approach, which is also known as the ‘solvability perspective’ or the ‘event characteristics’ perspective, postulates that it is rarely the case that investigators deliberately conduct their analyses in a different manner because of the sociodemographic characteristics of the victims of homicides (Gottfredson & Hindelang, 1979; Roberts, 2007). According to non-discretionary perspective, it is the specificities of the murder that matter most. It is the context of the murder and the crime scene that determine the possibility to solve the case rather than the victim’s economic status or ethnicity (Litwin, 2004; Litwin & Xu, 2007; Puckett & Lundman, 2003; Roberts, 2007). Consequently, the reasons for an investigation’s failure should be sought in the specific characteristics of the homicide, not in the sociodemographic characteristics of the victim. Studies adopting this perspective suggest that solely factors as the use of weapons other than firearms or the homicide taking place in a private residence determine a higher probability of homicide clearance (Klinger, 1997; Riedel, 2008).

Empirical research has provided mixed results with respect to discretionary perspective. Several scholars have found that murder cases with ethnic minority victims are more likely to be cleared than those with white victims (Litwin & Xu, 2007; Petersen, 2017; Regoeczi et al., 2000), whereas other scholars have observed the opposite (Litwin, 2004; Regoeczi et al., 2000; Roberts & Lyons, 2011) or no impact of the ethnicity on the result of the investigation (Addington, 2006; Jiao, 2007; Puckett & Lundman, 2003). Several studies have found that the victim’s gender has not any significant effect on

clearance (Addington, 2006; Puckett & Lundman, 2003), while other scholars have reported a higher likelihood of cases being cleared if they involve either female victims (Lee, 2005; Petersen, 2017; Roberts, 2007) or male victims (Jiao, 2007; Litwin & Xu, 2007; Roberts & Lyons, 2009, 2011). The literature indicates that homicides involving child victims are more likely to be cleared (Addington, 2006; Jiao, 2007; Roberts, 2007); while there is more debate on the relation between elderly people and homicide clearance (Bänziger & Killias, 2014; Liem et al., 2019, 2020; Regoeczi et al., 2020).

As regards the non-discretionary factors, the relevance of the location of the murder has received empirical support: homicides occurring indoors are more likely to be cleared, due to both the possibility of preserving evidence better and the likelihood that an indoor location entails a personal relationship between victim and perpetrator (Addington, 2006; Braga et al., 2019; Litwin, 2004; Litwin & Xu, 2007; Mouzos & Muller, 2001; Regoeczi et al., 2000). Homicides committed with firearms are usually considered to be less likely to be cleared than those committed with other weapons (Liem et al., 2020). This is because firearms involve less contact –and hence less evidence– between the opponents compared to other weapons, such as knives or blunt objects (Braga et al., 2019; Litwin, 2004; Litwin & Xu, 2007; Regoeczi et al., 2020; Roberts, 2007; Roberts & Lyons, 2011; Rydberg & Pizarro, 2014). In terms of population density of the area in which the homicide has occurred, the findings are mixed. Some studies evidence that clearance rates for homicides occurring in small cities and sparsely populated areas tend to be higher (Keel et al., 2009; Litwin, 2004; Mancik et al., 2018). On the other side, a few scholars have found that homicides in populated areas are more likely to be cleared than those in less populated areas (Wellford & Cronin, 1999). Finally, other studies have found no significant differences in clearance rates between urban and rural areas (e.g., Liem et al., 2019).

Current study

Especially in contexts other than the United States, much remains to be understood about the relationship between specific features of murders and of their investigations and the chances of murder cases being cleared (Brookman et al., 2019; Liem et al., 2019; Pizarro et al., 2020; Regoeczi et al., 2020; Vaughn, 2020). Related to this, the literature does not give a final answer on the explanatory power of the principal conceptual perspectives on the explanations of homicide clearances: the discretionary and the non-discretionary perspective (Campedelli, 2022; Liem et al., 2019). Accordingly, this study explores the association with the clearance of homicide cases in Italy of factors related to discretionary (Black, 1976; Paternoster, 1984; Peterson & Hagan, 1984) and to non-discretionary (Gottfredson & Hindelang, 1979; Klinger, 1997; Roberts, 2007) perspectives. In fact, we considered both sociodemographic characteristics of the victims –i.e., nationality, age, sex, socioeconomic status, previous convictions– which might relate to discretionary behavior by the police and factors pertaining to the murder’s situational domain –i.e., circumstance, weapon, location, urban setting, macro-area where the homicide took place, and number of victims in the same episode. Finally, to explore the relation between media pressure and police performance, we checked also for media coverage of the criminal episode. We performed our analysis on all murders occurring in Italy in 2014. Then, we replicated the analysis on the subsample of cases requiring greater investigative efforts: i.e., difficult-to-solve-cases.

Empirical Strategy

Outcome measure

As dependent variable, we focused on the outcome –i.e., identification of the murderer(s) or not– of all homicide investigations conducted in Italy in 2014. In the present study, the distinction between cleared and unsolved cases refers to cases in which at least one person has been arrested by the police, and the

arrest was subsequently validated by the Public Prosecutor's Office. The double-check process, based on the police arrest and the Prosecutor's subsequent validation, results in greater data reliability. Separately, we focused exclusively on the most challenging cases (i.e., difficult-to-solve-cases), as described in the next subsection. The original dataset comprised 458 victims pertaining to the entire population homicides that were either reported to the police or autonomously investigated by such authorities in Italy during 2014. Because of privacy issues, the information contained in the original dataset is partial and refers to the demographic characteristics of the victim(s) and the location in which the crime occurred up to the provincial level or, at the very most, the metropolitan level.

The database was then crosschecked and integrated with openly available information gathered from national and local media websites and, following the approach suggested by Hawk et al. (2021), from a multisite survey administered to detectives from the Carabinieri Corps operating throughout country, who participated in 98 of the difficult-to-solve investigations documented in the dataset. While surveys are usually exploited in single-site studies, the fact that the Carabinieri Corps operate as a single large police department across Italy made it possible to produce a multisite survey. After checks, 3 homicides proved to be natural deaths (0.65% of all homicides) and 3 were road accidents (0.65%). In 9 cases (1.97%), it was impossible to retrieve information on the investigation and its outcome. We removed these homicides from the analysis, resulting in 443 victims. In turn, these 443 identified victims corresponded to 404 different cases because of 31 cases concerning multiple victims –373 single homicides (92.33% of total cases), 23 double homicides (5.69%), 8 triple homicides (1.98%). The homicide was cleared in 350 (79.0%) cases and remained unsolved in the remaining 93 (21.0%).

Easy and difficult to-solve cases

To take account of the nature of the homicide cases themselves, for each statistical approach, we ran two sets of analyses: one focusing on all homicide cases and the second considering only those cases

which required a significant amount of investigative work (i.e., difficult-to-solve-cases). Isolating difficult-to-solve cases is important because the complexity of the investigation may influence both its outcome and discretionary actions by the police. In fact, discretion is less likely to emerge in easy-to-solve cases, which require less efforts by the investigators (Riedel 2008). In the empirical literature, there is no consensus regarding the criteria to use in determining whether a case is difficult or easy to solve. In Italy, investigative units deal only with homicide cases which require a certain level of investigative action. Exploiting this feature of the Italian policing system, we primarily classified homicide cases as difficult-to-solve-cases when primary and auxiliary investigative units directly handled the investigation, due to the puzzling circumstances of the homicide case.¹ Conversely, we labelled as easy-to-solve those simple cases (e.g., domestic homicides, when the perpetrator is known from the very beginning), which were exclusively managed by non-investigative units, such as patrol units or station command personnel. We labelled as easy-to-solve cases also those homicides involving surrendered perpetrators and those in which the perpetrator committed or attempted a homicide/suicide, irrespectively of the type of unit that dealt with the case. Through the application of these criteria, we identified 212 difficult-to-solve cases (47.9% of the total; clearance rate 56.1%) and 231 easy-to-solve cases (52.1% of the total; clearance rate 100.0%). The ratio between easy-to-solve and difficult-to-solve cases is consistent with the previous literature, in which easy-to-solve cases commonly range between half and two-thirds of all homicides (e.g., Alderden & Lavery, 2007).

¹ 'Primary investigative units' are units which operate at provincial level, are based in the provincial capital, are specialized in homicide investigations, and are in charge of major investigations such as homicides, robbery, organized crime, fugitives, and so forth. 'Auxiliary units' operate at local level and, although they are typically responsible for dealing with crimes such as burglary, robbery, extortion, drug dealing, attempted homicide, they often perform murder investigations as well.

Selection and operationalization of homicide case features

Discretionary: individual domain factors

As first, we investigated the role of nationality, age, and gender of the victim. Unlike in North America, in Italy, it makes more sense to examine *nationality* rather than race, because discrimination more broadly concerns migrants (Zanfrini, 2019). Bänziger and Killias (2014) used the same operationalization in their study on homicide clearance in Switzerland. As previously discussed, homicides involving victims of a younger *age* are usually cleared more frequently, while there is more debate on the clearance of those cases involving victims aged over 65. In consideration of the frequency of homicide victimization by age and the limited relevance of youth gangs in Italy, we then constructed a categorical variable, which separates victims 25 years-old or younger, aged between 25 and 65 years, and older than 65. From a discretionary perspective, homicides with female victims are more likely to be solved than ones with male victims (Riedel, 2008). Data on these victims' characteristics were drawn from the original dataset produced by the Ministry of the Interior.

One of the central tenets of the discretionary perspective is that the police base their level of commitment upon the *socioeconomic status* (SES) of either the victim or the neighborhood in which the homicide has occurred. Therefore, disadvantaged victims receive 'less law' and consequently record lower levels of homicide clearance. To test this hypothesis, we relied on the information emerging from the above-mentioned survey. The respondents were asked to assess the socioeconomic status of the victims in the homicide cases that they had investigated by classifying them into: 'wealthy', 'middle-class', 'working-class', 'indigent'. For the 104 difficult-to-solve-cases not covered by the survey, the scanning of open sources yielded information related to victims' socioeconomic status (e.g., direct description of the victims' socioeconomic status, references to their job or to the neighborhood where they lived), which we classified according to the same four SES categories. We assessed the degree of

convergence between the two sources of information by conducting the media content analysis on a sample of victims for which we had the results of the survey, and we observed high consistency in the assessment of the socioeconomic status. Finally, we grouped ‘wealthy’ (8 cases, 3.8% of the total; Clearance Rate (CR) 62.5%) and ‘middle-class’ (91, 42.9%; CR 62.6%) and, separately, ‘working-class’ (98, 46.2%; CR 50.0%) and ‘indigent’ (15, 7.1%; CR 53.3%) to have a larger number of observations for category. Adopting the same approach (i.e., survey and content analysis of news) for difficult-to-solve cases, we evaluated also whether the victims had *previous criminal records* (108, 50.9%) or not (104, 49.1%).

Besides sociodemographic factors, we controlled for the *media coverage* of difficult-to-solve cases, considering that media pressure may induce investigators to put extra efforts into their investigations. Three categories were used to classify homicide cases: low media coverage (108, 50.9%), i.e. whether information on the homicide case was retrievable only on local news outlets; medium media coverage (71, 33.5%), i.e. whether national newspapers covered the homicide; high coverage (33, 15.6%), i.e. whether multiple national newspapers and agencies paid attention to the homicide.

Non-discretionary: event and context characteristics

The micro-level information on victims and events provided in the original dataset made it possible, by means of open source analysis, to match the 443 victims with the respective homicide case and investigation, hence increasing the number of variables included in our analysis. With respect to the *weapon* used in homicide, we distinguished between ‘firearms’ and ‘non-firearms’ as previously done by scholars such as Litwun (2004) or Regoeczi and colleagues (2020). Another key factor impacting on homicide clearance according to non-discretionary perspective is the *circumstance* in which the homicide emerges and, consequently, the motivation behind it (Alderden & Lavery, 2007; Litwin, 2004; Litwin & Xu, 2007). Regarding the circumstance, in accordance with the literature (e.g., Alderden & Lavery,

2007), we classified the incidents into three categories: ‘criminal-instrumental’, comprising criminal liquidation, other criminal motives, drugs- or sex-related homicides, or more trivial reasons; ‘intimate-expressive’, comprising both domestic and acquaintance disputes; ‘unknown’ motive, meaning that the investigators do not know the reasons for the homicide. The crime scene *location* was classified according to the difference between homicides that occurred ‘indoors’ or ‘outdoors’. We also considered whether the homicide took place in *urban setting* or not differentiating between homicides in a ‘city’ (the capital cities of the 110 Italian provinces existing in 2014) and in the ‘countryside’ (any other municipality). At the same time, most mafia homicides occur in Southern Italy, which required us to include controls for the different *macro-areas* in which homicides and investigations took place: Center-North and South. Finally, the number of victims was also considered, since previous studies had reported that cases involving multiple victims are more likely to be cleared than those involving one victim because they may call for closer attention (Mouzos & Muller, 2001). In our analysis, the spectrum ranged from one to three victims.

Statistical approach

To determine if there are any statistically significant associations between each of the categorical variables present in the data and whether the case was cleared or unsolved, we first performed Pearson’s (1900) χ^2 tests of independence. As second, we analyzed if and how much sociodemographic characteristics and specificities of the case considered were correlated with the clearance rate of the homicide cases in a multivariate fashion. Because of the dichotomous outcome variable (i.e., unsolved and cleared cases), we exploited Logistic Regression (LR) models to perform this analysis. Binary response logistic regression is based on the Bernoulli probability distribution, and it does not require independent variables to be linearly related; nor does it require homoscedasticity, which makes it a quite flexible instrument for statistical analysis (Hosmer et al., 2013).

Finally, we implemented an exploratory inductive approach based on machine learning algorithms. A random forest classification is constructed by combining multiple classification trees that together predict the response variable (see Breiman 2001). In turn, the recursive binary partitioning that creates the classification trees is a non-parametric technique that helps one explore the structure of a set of data (Breiman et al., 1984; Ripley, 1996). Then we used two evaluation procedures to assess the efficacy of discretionary and non-discretionary factors in predicting homicide clearance. First, we used Area Under the Curve (AUC) scores to compare the predictive performance of a fully specified random forest which includes both discretionary and non-discretionary factors with the predictive performances of both a ‘Discretionary’ specification and a ‘Non-discretionary’ specification. The ‘Discretionary’ specification of the forest considers only factors potentially inducing discretionary behaviors. The ‘Non-discretionary’ specification includes exclusively factors unrelated to discretionary behaviors. Operatively, we used cross-validation procedures to estimate out-of-sample accuracy thus evaluating the predictive efficacy of our model specifications. We trained our models on a random training-test split (.70 of observations) and we used the remaining observations (.30 of the total) as a hold-out test. We used the R *partykit* library by Hothorn and Zeileis (2015) to perform the random forest. We set the number of trees to 750 and the maximum tree depth to infinite as the statistical approach adopted by *partykit* ensures that the right-sized tree is grown without additional post-pruning or cross-validation (Hothorn et al., 2021).

Second, we evaluated the contribution of each single factor to the random forest capacity to correctly predict clearances of cases. To conduct this analysis, we estimated the permutation accuracy importance, which highlights how the accuracy of predictions reduces when a variable is removed from the model. In doing so, permutation accuracy importance considers both the impact of each predictor variable individually as well as its impact in multivariate interactions with other predictor variables.

Indeed, the importance of each variable is computed by permuting within a grid defined by the covariates that are associated to the variable of interest (Figure 1) (see Strobl et al. 2007, 2009). Such an approach based on predictive accuracy and variable importance cannot produce causal claims. However, it allows for gain insights into the interaction among attributes and further disentangles what factors discriminate cleared and unsolved homicides without making difficult assumptions concerning parameter estimates and standard error accuracy.

Figure 1. Correlation matrix of predictors



Note: ***, **, * indicate significance of the correlation at 0.1%, 1% and 5%.

Results

Pearson's χ^2 test

When considering all homicides, Pearson's χ^2 test identifies statistically significant differences in the distributions of cleared and unsolved cases with respect to age ($p < .01$) and sex ($p < .001$) of the victim (Table 1). Unsolved cases were more likely to involve firearms that not ($p < .001$), to have a criminal-instrumental motivation ($p < .001$), and to be committed in the South of Italy rather than in the Center-North ($p < .001$). Focusing exclusively on difficult-to-solve cases, the relationship between the media coverage of the case and its clearance is also significant ($p < .001$); 93.9% of cases with a high media coverage had been solved, while the shares were 60.6% and 41.7% for cases with a medium or low coverage, respectively. The Pearson's χ^2 test identifies no significant relationship between victim's nationality, socioeconomic status, criminal records, and the clearance of both the easy-to-solve and the difficult-to-solve-cases. Moreover, we observe no significant relationship between the location (i.e., indoor-outdoor) of the difficult-to-solve homicides and whether the cases were solved. Similarly, the homicide occurring in a city or in the countryside does not emerge as being associated to the outcome of the investigations.

Table 1. Sample characteristics and Pearson's χ^2

			<i>All cases</i>			<i>Difficult-to-solve-cases only</i>		
			Cleared	Unsolved	Total	Cleared	Unsolved	Total
<i>Nationality/ethnicity</i>			$p=0.210$ (n.s.) $V=0.060$			$p=0.267$ (n.s.) $V=0.076$		
Italian	N.		277	79	356	94	79	173
	%		77.81	22.19		54.34	45.66	
Noncitizens	N.		73	14	87	25	14	39
	%		83.91	16.09		64.10	35.90	
<i>Age</i>			$p=0.008$ (**) $V=0.147$			$p=0.042$ (*) $V=0.173$		
<25	N.		59	6	65	20	6	26
	%		90.77	9.23		76.92	23.08	
26-65	N.		218	73	291	77	73	150
	%		74.91	25.09		51.33	48.67	
>65	N.		73	14	87	22	14	36
	%							

		%	83.91	16.09		61.11	38.89	
<u>Sex</u>			$p=0.000 (***)$ $V=0.289$			$p=0.001 (***)$ $V=0.231$		
Female	N.		134	5	139	26	5	31
	%		96.40	3.60		83.87	16.13	
Male	N.		216	88	304	93	88	183
	%		71.05	28.95		51.35	48.62	
<u>SES</u>			Not included			$p=0.074 (n.s.)$ $V=0.123$		
Middle-class & Wealthy	N.					62	37	99
	%					62.63	37.37	
Working-class & Indigents	N.					57	56	113
	%					50.44	49.56	
<u>Criminal record</u>			Not included			$p=0.120 (n.s.)$ $V=0.107$		
Yes	N.					55	53	108
	%					50.93	49.07	
No	N.					64	40	104
	%					61.54	38.46	
<u>Media coverage</u>			Not included			$p=0.000 (***)$ $V=0.369$		
High	N.					31	2	33
	%					93.94	6.06	
Medium	N.					43	28	71
	%					60.56	39.44	
Low	N.					45	63	108
	%					41.67	58.33	
<u>Number of Victims</u>			$p=0.389 (n.s.)$ $V=0.041$			$p=0.771 (n.s.)$ $V=0.020$		
Multiple	N.		58	12	70	17	12	29
	%		82.86	17.14		58.62	41.38	
Single	N.		292	81	373	102	81	183
	%		78.28	21.72		55.74	44.26	
<u>Weapon</u>			$p=0.000 (***)$ $V=0.382$			$p=0.000 (***)$ $V=0.441$		
Firearm	N.		116	74	190	42	74	116
	%		61.05	38.95		36.21	63.79	
Non-firearm	N.		234	19	253	77	19	96
	%		92.49	7.51		80.21	19.79	
<u>Circumstance</u>			$p=0.000 (***)$ $V=0.765$			$p=0.000 (***)$ $V=0.699$		
Criminal-instrumental	N.		68	49	117	50	49	99
	%		58.12	41.88		50.51	49.49	
Intimate-expressive	N.		281	1	282	68	1	69
	%		99.65	0.35		98.55	1.45	
Unknown	N.		1	43	44	1	43	44
	%		2.27	97.73		2.27	97.73	
<u>Location</u>			$p=0.000 (***)$ $V=0.279$			$p=0.132 (n.s.)$ $V=0.104$		
Indoors	N.		199	21	220	38	21	59
	%		90.45	9.55		64.41	35.59	
Outdoors	N.		151	72	223	81	72	153
	%		67.71	32.29		52.94	47.06	

<i>Setting</i>		<i>p=0.260 (n.s.) V=0.054</i>			<i>p=0.420 (n.s.) V=0.055</i>		
Countryside	N.	247	60	307	83	60	143
	%	80.46	19.54		58.04	41.96	
City	N.	103	33	136	36	33	69
	%	75.54	24.26		52.17	47.83	
<i>Macro-area</i>		<i>p=0.000 (***) V=0.304</i>			<i>p=0.000 (***) V=0.340</i>		
Center-North	N.	227	26	253	74	26	100
	%	89.72	10.28		74.00	26.00	
South	N.	123	67	190	45	67	112
	%	64.74	35.26		40.18	59.82	
Total		N.	350	93	443	119	93
		%	79.01	20.99	100.0	56.13	43.87
						100.0	

Note: For each category, the first row has frequencies, and the second row has percentages. ***, **, *, n.s. indicate significance of the Pearson's χ^2 test at 0.1%, 1%, 5%, > 5%. Cramér's V (1946) (i.e., *V*) provides a measure of the degree of association between the nominal variable considered and the outcome of the investigations.

Logistic Regression

We further investigated homicide clearance by means of two sets of logistic regression models: the first considers all homicide cases (All-LR, Table 2); the second focuses only on difficult-to-solve-cases (Difficult-LR, Table 3). We first use as regressors nationality, age, gender: in this model, whether the victim was Italian or not is not correlated with the probability of the homicide case being cleared; homicide cases in which the victim is an adult (25-65) are significantly less likely to be solved than cases in which the victim is less than 25 years old; homicide cases with male victims are less likely of being cleared relative to homicides with a female victim (OR .09, $p < .001$ in model All-LR.1). In model All-LR.2, we added to our regression the number of victims involved in the same case and the weapon used to commit the homicide. The use of firearms is significantly and negatively correlated with the homicide being solved (OR .17, $p < .001$) holding other things constant. Cases involving a single victim, instead, do not differ from homicides with multiple victims in terms of clearance rate, other considered factors held equal.

We then expanded the list of regressors to include the circumstance in which the homicide occurred. The circumstance emerges as the single factor most strictly related to the probability of a

homicide case being cleared. Criminal-instrumental homicides and homicides with an unknown motive are significantly less likely to be cleared than homicides with an intimate or an expressive motive (All-LR.3). Finally, we extended the set of regressors by controlling for the location, the rural or urban setting, and the macro-area in which the homicide was committed. Of these variables, only the macro-area is significantly correlated to the clearance rate. In the South, homicides are less likely to be cleared than in the Center-North (OR .03, β -0.891, $p < .01$ in model All-LR.4). In this richer regression set, macro-area, sex, and circumstances are the only predictors significantly correlated to the clearance rate. Homicide circumstances have the largest effect size among the considered predictors: the absolute values of the standardized β coefficients of criminal-instrumental and unknown circumstances are 1.923 and 3.454 respectively.

Table 2. All cases, logistic regressions

		DV: Homicide Clearance (0;1)			
		All-LR.1	All-LR.2	All-LR.3	All-LR.4
<u>Nationality and ethnicity</u>					
	Noncitizens	1.47 (.58) [.332] {0.152}	0.87 (.38) [.739] {-0.057}	2.36 (1.93) [.295] {0.341}	1.02 (1.23) [.985] {0.009}
<u>Age</u>					
	26-65	0.29** (.12) [.004] {-0.593}	0.41 (.19) [.058] {-0.425}	0.51 (.27) [.200] {-0.316}	0.61 (.31) [.337] {-0.235}
	>65	0.38 (.19) [.056] {-0.382}	0.29* (.15) [.016] {-0.491}	0.55 (.52) [.525] {-0.241}	.61 (.59) [.609] {-0.200}
<u>Sex</u>					
	Male	0.09*** (.04) [.000] {-1.116}	0.14*** (.07) [.000] {-0.926}	0.17 (.16) [.053] {-0.828}	0.12* (.10) [.011] {-0.991}
<u>Number of victims</u>					
	Single		0.71 (.36) [.501] {-0.126}	0.45 (.28) [.206] {-0.294}	0.49 (.34) [.300] {-0.258}
<u>Weapon</u>					
	Firearm		0.17*** (.05) [.000] {-0.891}	0.13** (.90) [.003] {-0.998}	0.34 (.28) [.192] {-0.532}
<u>Circumstance</u>					
	Criminal-instrumental			0.01***	0.01***

			(.01) [.000] {-1.996} 0.00***	(.01) [.000] {-1.923} 0.00***
Unknown			(.00) [.000] {-3.086}	(.00) [.000] {-3.454}
<u>Location</u>				
Indoors				0.86 (.62) [.832] {-0.077}
<u>Setting</u>				
City				0.75 (.64) [.737] {-0.133}
<u>Macro-area</u>				
South				0.03** (.04) [.005] {-1.714}
Constant: included				
<i>N</i>	443	443	443	443
Parameters	8	12	15	21
Log pseudolikelihood	-198.9	-179.6	-74.6	-64.0
Likelihood-ratio χ^2		38.57 (.000)	209.91 (.000)	21.16 (.000)
Count R^2	0.79	0.79	0.92	0.93
Adj. Count R^2	0.00	0.02	0.62	0.68
McFadden's Pseudo- R^2	0.13	0.21	0.67	0.72
Adj. McFadden's R^2	0.09	0.16	0.61	0.63
Bayesian crit. (BIC)	428.2	401.8	204.1	201.2
Degrees of freedom	4	6	8	11
Pearson χ^2	3.32	42.43	44.64	84.02
Prob > χ^2	.650	.066	.883	1.00
AUC	0.72	0.81	0.97	0.98

Note: Odds ratios; exponentiated robust standard errors clustered at the municipal level in parentheses (); p -values in square brackets []; x-standardized coefficient (i.e., β) in curly braces {}. The Likelihood-ratio χ^2 statistic is used to test the null hypothesis of identical fit for any pair of subsequent models; p -value of the test is in parentheses. The Count R^2 , the adjusted Count R^2 and the pseudo- R^2 provide indications on the proportion of the variance in the dependent variable that is predictable from the independent variables. The BIC value provides a measure of the relative quality of the model. The Pearson goodness-of-fit test assesses whether the observed event rates match expected event rates in subgroups of the model population (i.e., the deciles of fitted risk values). The AUC is a perfect performance metric for the ROC curve which summarizes the model's performance by evaluating the trade-offs between true positive rate (sensitivity) and false positive rate (1 - specificity). Reference categories are: Nationality (Italian); Age (0-25); Sex (Female); Number of Victims (Multiple); Weapon (Non-firearms); Circumstance (Intimate-expressive); Location (Outdoors); Setting (Countryside); Macro-area (Center-North). ***, **, * indicate significance at 0.1%, 1% and 5%.

We then analyzed difficult-to-solve-cases only (Table 3). In this analysis, we explore the role of the socioeconomic status, the victim criminal record, and the media coverage of the homicide controlling for the factors emerged as significant from the Pearson's χ^2 test (Table 1). The correlations between homicide clearance as well as the socioeconomic status and the victim criminal record are not significant

both when considered separately (Difficult-LR.2 and Difficult-LR.3) as well as when observed in the most complete model (Difficult-LR.5). While a high media-coverage correlates with a higher clearance rate, a low media-coverage is not correlated with low clearance rate, all else being equal. In the models considering difficult-to-solve-cases too, the macro-area (Difficult-LR.5), the motive for the homicide (Difficult-LR.4 and Difficult-LR.5), and the sex (all models but Difficult-LR.4) display a significant correlation with the clearance rate of homicides. Use of firearms and age lose significance in the most complete model (Difficult-LR.5). Differently from the analysis of the full sample, the age of the victim is significantly correlated to clearance in most regressions (all models but the richest one).

Table 3. Difficult-to-solve-cases, logistic regression

		DV: Homicide Clearance (0;1)				
		Diff.-LR.1	Diff.-LR.2	Diff.-LR.3	Diff.-LR.4	Diff.-LR.5
<u>Age</u>						
	26-65	0.26* (.14) [.011] {-0.619}	0.25* (.14) [.011] {-0.637}	0.22** (.13) [.008] {-0.694}	0.32* (.18) [.047] {-0.515}	0.41 (.22) [.102] {-0.410}
	>65	0.25 (.26) [.184] {-0.521}	0.23 (.24) [.167] {-0.552}	0.28 (.27) [.194] {-0.508}	0.36 (.40) [.355] {-0.380}	0.86 (.97) [.862] {-0.078}
<u>Sex</u>						
	Male	0.13* (.13) [.040] {-0.733}	0.13* (.13) [.046] {-0.726}	0.13* (.12) [.037] {-0.773}	0.31 (.30) [.229] {-0.417}	0.09* (.09) [.017] {-0.876}
<u>Weapon</u>						
	Firearm	0.13** (.09) [.003] {-1.035}	0.12** (.09) [.003] {-1.061}	0.09** (.08) [.002] {-1.128}	0.11** (.08) [.002] {-1.100}	0.14 (.15) [.059] {-0.968}
<u>Circumstance</u>						
	Criminal-instrumental	0.03*** (.03) [.000] {-1.734}	0.03*** (.03) [.000] {-1.752}	0.02*** (.02) [.000] {-1.824}	0.03*** (.02) [.000] {-1.793}	0.02*** (.02) [.000] {-1.840}
	Unknown	0.00*** (.00) [.000] {-3.579}	0.00*** (.00) [.000] {-3.578}	0.00*** (.00) [.000] {-3.646}	0.00*** (.00) [.000] {-3.473}	0.00*** (.00) [.000] {-4.379}
<u>Macro-area</u>						
	South					0.02** (.03) [.008] {-1.859}
<u>SES</u>						
	Working-class & Indigents		0.79 (.37) [.608] {-0.121}			0.42 (.28) [.188] {-0.435}

<i>Criminal record</i>					
Yes			2.01 (.78) [.328] {0.238}		2.91 (1.34) [.070] {0.473}
<i>Media coverage</i>					
High			9.77** (7.78) [.004] {0.818}		10.42** (10.63) [.005] {0.893}
Low			0.69 (.41) [.529] {-0.186}		1.10 (.73) [.874] {0.052}
Constant: included					
<i>N</i>	212	212	212	212	212
Parameters	11	13	13	14	20
Log pseudolikelihood	-64.7	-64.5	-64.0	-60.1	-49.0
Likelihood-ratio χ^2		0.29(.591)	0.66(.415)	9.15(.010)	31.00(.000)
Count R ²	0.86	0.86	0.86	0.87	0.90
Adj. Count R ²	0.68	0.68	0.68	0.71	0.77
McFadden's Pseudo-R ²	0.56	0.56	0.56	0.59	0.66
Adj. McFadden's R ²	0.48	0.47	0.47	0.49	0.53
Bayesian crit. (BIC)	-947.3	-936.9	-938.0	-940.4	-930.4
Degrees of freedom	6	7	7	8	11
Pearson χ^2	12.03	32.11	16.86	40.51	58.12
Prob > χ^2	.799	.654	.987	.448	1.00
AUC	0.93	0.93	0.94	0.94	0.96

Note: Odds ratios; exponentiated robust standard errors clustered at the municipal level in parentheses (); *p*-values in square brackets []; x-standardized coefficient (i.e., β) in curly braces {}. The Likelihood-ratio χ^2 statistic is used to test the null hypothesis that the simplest model (Diff.-LR.1) provides as good a fit for the data as the considered model (Diff.-LR.2 to Diff.-LR.5); *p*-value of the test is in parentheses. Reference categories are: Nationality and ethnicity (Italian); Age (0-25); Sex (Female); SES (Middle-class & Wealthy); Criminal record (No); Number of Victims (Multiple); Media Coverage (Medium); Weapon (Non-firearms); Circumstance (Intimate-expressive); Location (Outdoor); Setting (Countryside); Macro-area (Center-North). ***, **, * indicate significance at 0.1%, 1% and 5%.

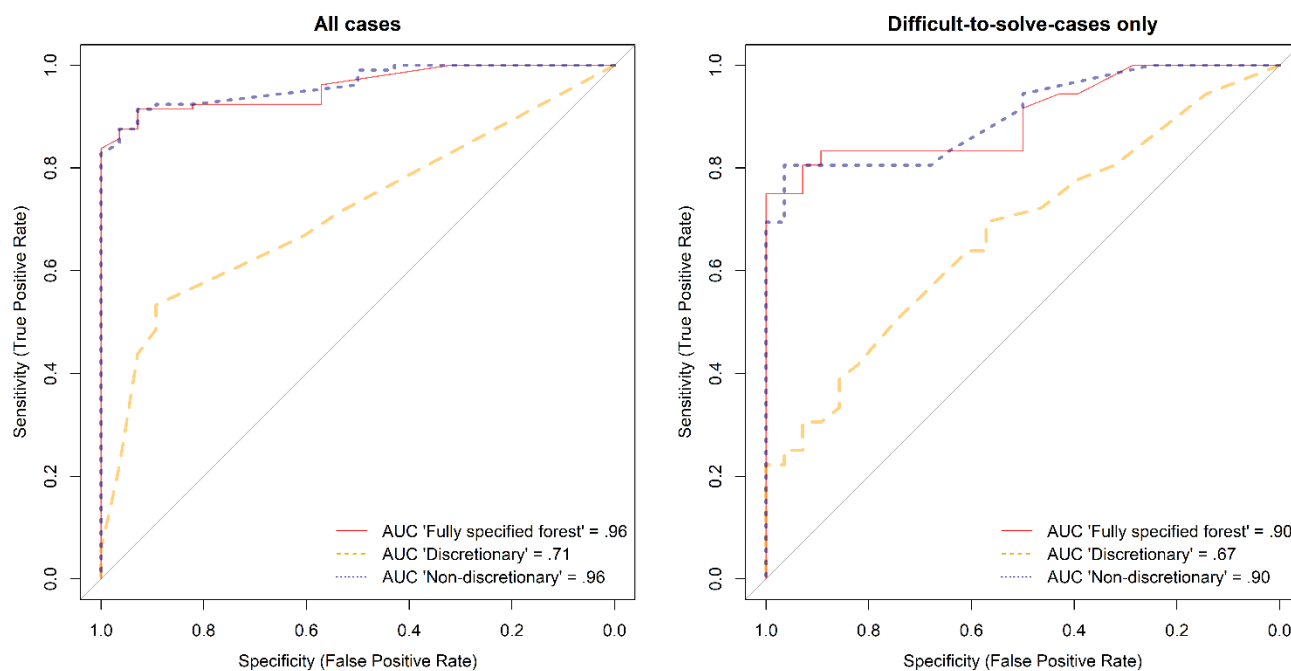
Random Forest Exploratory Models

Out-of-Sample Accuracy Estimation

The fully specified random forest has an AUC of .96, when considering all homicides (Figure 2, left). The random forest specification considering exclusively non-discretionary factors has also an AUC of .96. The ‘discretionary’ specification has an AUC of .71. Therefore, the performance of the ‘discretionary’ specification is fifteen-point lower than the performance of the ‘non-discretionary’ specification. The capacity of the random forest to predict homicide clearance of difficult-to-solve-cases is lower than it is for all homicides, even though it exploits information on SES, previous convictions, media coverage, in addition to the predictors used also to classify the full sample of cases. In fact, the

fully specified random forest has an AUC of .90 (Figure 2, right). With respect to difficult-to-solve-cases too, the relative performance of the ‘non-discretionary’ specification is remarkably similar to the one considering all predictors (AUC .90). Finally, the ‘discretionary’ specification has an AUC of .67. Therefore, the analysis of the out-of-sample accuracy estimation indicates that the accuracy of the fully specified random forest is mostly driven by non-discretionary factors. A specification which considers the event characteristics –i.e., ‘non-discretionary’– performs as good as a specification that considers both event characteristics and individual domain factors. At the same time, the AUCs of the specifications considering exclusively ‘discretionary’ factors have quite poor predictive accuracies.

Figure 2. ROC curves for random forest specifications

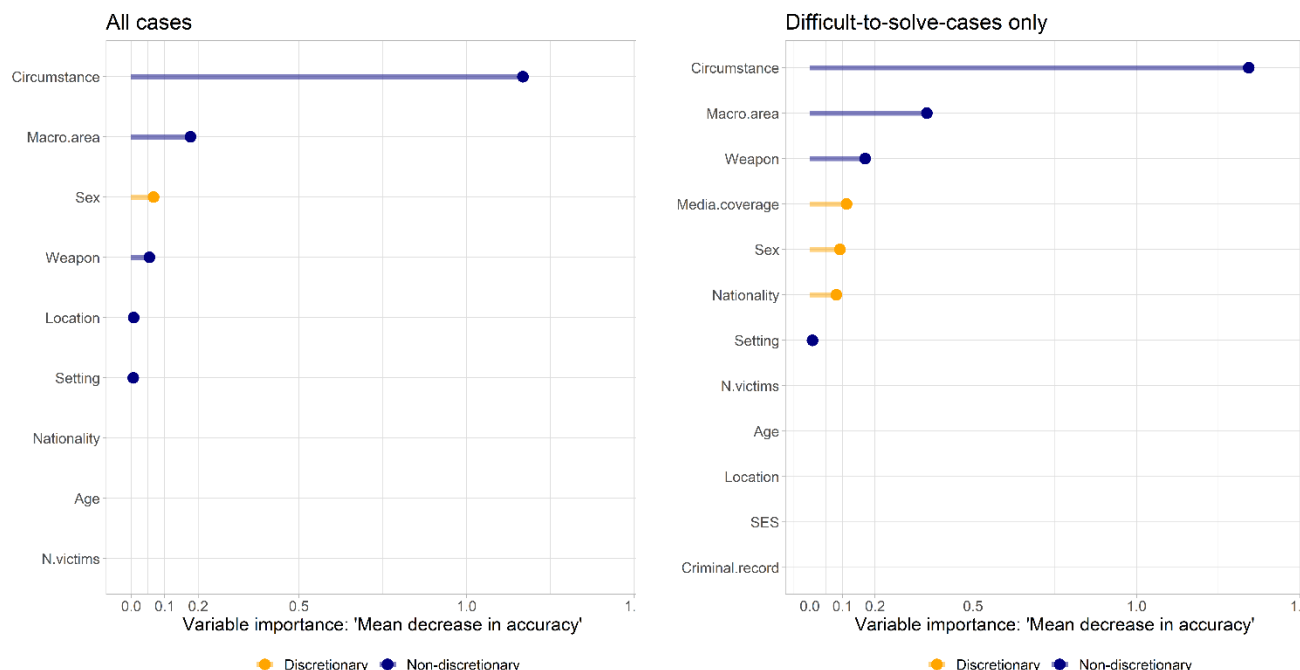


Note: ROC curves plot sensitivity versus specificity, i.e., accuracy in identifying the majority class. The AUC scores provide measurements of the performance of classifications, which range from random guessing (AUC = .50) to perfect prediction (AUC = 1.00). Three specifications have been considered. ‘Fully specified forest’ includes all predictors. ‘Discretionary’ includes: age, nationality, and sex of the victim, when considering all cases. When considering difficult-to-solve-cases, ‘Discretionary’ includes also socioeconomic status and criminal records of the victim as well as the media coverage of the case. ‘Non-discretionary’ includes: circumstance, weapon, location, setting, macro-area, and number of victims.

Variable Importance

As further step, we evaluated the impact of each feature associated with a homicide case on the predicted outcome of the investigation by estimating the average decrease in overall predictive accuracy when randomly permuting the original values of a predictor. The circumstance of the homicide emerges as the factor having the greatest impact on the capacity of the models to predict clearance of the homicides (Figure 3). The macro-area where the homicide occurred is the factor causing the second largest decrease in accuracy when left out, both when considering all cases and when focusing on difficult-to-solve-cases. The impact on prediction performances of all other factor is lower, although with differences between factor and factor. When considering the full sample of homicides, sex of the victim and employed weapon have the third and fourth strongest impact on prediction accuracy; all other predictors have a minimum impact on the performances of the random forests. When considering exclusively difficult-to-solve-cases, weapon, followed by media coverage, sex and nationality of the victim contribute to model prediction accuracy. All other factors have a minimum impact on prediction accuracy. The lack of importance for many discretionary factors explains why the ‘discretionary’ specification performed poorly compared to the specification based on non-discretionary factors.

Figure 3. Variable importance



Note: The variable importance is measured as the standardized difference in the out-of-sample prediction accuracy before and after permuting 30 times each variable, averaged over all trees forming the forest (i.e., 750).

Discussion

Although conflicting evidence has been found with respect to discretionary action by criminal investigators, there is nonetheless a wealth of studies that have shown the association between extra-legal factors and clearance, specifically those pertaining to victims' age, gender, and ethnicity (e.g., Addington, 2006; Alderden & Lavery, 2007; Campedelli, 2022; Jiao, 2007; Lee, 2005; Litwin & Xu, 2007). At the same time, several studies have reported a significant correlation between the homicide clearance rate and factors related to the criminal event and its context such as the weapon used, the time and location of the homicide, and the motivations behind the homicide (e.g., Alderden & Lavery, 2007; Carter & Carter, 2016; Keel et al., 2009; Rydberg & Pizarro, 2014). The empirical results presented in this paper do not provide support for the role of most discretionary factors in the clearance of homicide cases in Italy.

Overall, the performed analyses suggest that most features possibly related to discretionary behavior are likely to be of secondary importance in determining whether a homicide case is cleared. Indeed, the logistic regressions did not identify any relation between most factors pertaining to the victim's individual domain and homicide clearance. The analyses provide no evidence that nationality, socioeconomic status, and previous criminal records of the victim are significant predictors of the outcome of homicide investigations. The sex of the victim is the only victim's characteristic which emerges as related to homicide clearance in all analyses: police are more effective in solving cases in which the victim is a woman. Out-of-sample accuracy analyses indicate that factors connected to non-discretionary perspectives contribute more than factors related to discretionary perspectives to the prediction accuracy of performed random forests. Coherently, the variable-importance analyses underlined the prominence of event and context characteristics in shaping prediction performances of random forests. Whether the homicide has criminal-instrumental motivations or intimate-expressive ones emerged as the factor most strongly connected to homicide clearance in Italy.

The nationality of the victim emerges as a relevant predictor of clearance only in the analysis of variable importance and only for difficult-to-solve-cases; the effect size of this relation was the smallest among statistically significant ones. In fact, the clearance rate for immigrants was higher than it was for Italians both in the general sample of murder cases and when only difficult-to-solve-cases were considered, but these differences were not statistically significant. With respect to Italy, this result does not support the conception, predicated on the discretionary perspective, that immigrants receive less legal protection or, more specifically, less commitment from the police (Black, 1976). Analogous results had been observed also in Finland and Switzerland (Bänziger & Killias, 2014; Liem et al., 2019).

Clearance was overall lower for victims aged 26 to 65. However, multivariate models in which we controlled for the macro-area where the homicide took place and analyses based on random forests, which

take into consideration the context of the specific case, did not identify age as a relevant predictor. None analyses of the socioeconomic status of the victim indicates that cases involving less affluent victims receive less attention from the Italian police. Future studies, based on larger samples, could investigate whether the victim's socioeconomic condition proves to be a significant predictor of a case being cleared, using more sensitive classifications of the socioeconomic status. By contrast, most of our analyses suggest that the sex of the victim is related to the chance of the homicide being cleared. Murders involving male victims are disproportionally less likely to be cleared than cases in which the victim is female. The relevance of sex or gender of the victim confirms the findings of previous studies conducted in other countries (e.g., Campedelli, 2022; Lee, 2005; Petersen, 2017; Roberts, 2007).

The fact that the removal of media coverage from the set of predictors of the clearance of difficult-to-solve-cases reduces prediction accuracy of random forests may indicate that police forces exercise discretion when investigating homicides based on extra-legal characteristics. Nevertheless, the modulation of the effort, in this case, would not be driven by the victims' socioeconomic status and demographics, but instead by the pressure of the media and, in turn, of public opinion. For instance, in 2010, the investigators in charge of identifying the murderer of a 13-year-old girl were able to examine as many as 18,000 DNA samples taken from relatives of the deceased and from many thousands of other males who were either local or known to have been in the area around the time of the victim's disappearance (Graversen et al., 2019). This unprecedented effort was, at least partially, related to the overwhelming public exposure of the case. Additionally, the relation between media coverage and homicide clearance has to be interpreted with prudence as it might be that public attention escalates when an investigation is successful.

Very much in line with non-discretionary perspectives as well as with the victim lifestyle perspective, all performed analyses stress the differences between homicides taking place in the

underworld and those committed within intimate disputes or as a reaction to stress. In particular, the variable importance analysis and the standardized regression coefficients identify the circumstances of the murder as the most relevant classification criterion both when considering all murders and when considering difficult-to-solve-cases. The importance of the motivation confirms that homicides are less likely to be cleared when there is a concomitant felony, or when the homicide is drug-related or organized crime-related, compared with homicides where a general altercation constitutes the main motive (e.g., Campedelli 2022; Litwin and Xu 2007; Regoeczi, Kennedy, and Silverman 2000). At the same time, the data suggest that homicide cases are more difficult to solve when they involve victims who engage in criminal activities because these cases are objectively harder to solve, rather than to a devaluation of criminal victims. Indeed, victims' previous criminal records are not correlated with the clearance rate.

Most of the literature has found that homicides committed in private locations are the most likely to be cleared (Addington, 2006; Litwin & Xu, 2007; Mouzos & Muller, 2001). In Italy too, the clearance rate for homicides committed indoors is significantly higher than for homicides committed outdoors (Table 1). Yet, the location of the homicide emerges as correlated to the outcome of the investigation neither when considering only difficult-to-solve-cases nor in any multivariate analysis. This result supports the contention that 'indoor' cases are easier to solve because the perpetrator is very often the (estranged) intimate partner or another family member (Litwin & Xu, 2007). This interpretation is further supported by the fact that, in our regression models which include a control for the murder's circumstance, the location of the homicide is not a significant predictor of the case's solution.

Few studies find that homicides in densely populated areas are more likely to be cleared than those in less populated areas, which can be attributed to the greater presence of witnesses in densely populated areas and to their willingness to share information (Wellford & Cronin, 1999). Several American studies, conversely, find that homicides occurring in a small town or a sparsely populated area are more likely to

be cleared (e.g., Keel et al., 2009; Litwin, 2004; Mancik et al., 2018). We find no evidence of significant differences in clearance rates between major cities and the countryside. Also in this regard, the Italian results show similarities with previous findings referring to other European countries (e.g., Liem et al. 2019). Further studies could investigate which structural or social characteristics of European cities and rural areas may play a role in determining this result.

Finally, all the analyses performed indicate a large and significant difference in the clearance rate between the Center-North of Italy and the South. Criminal-instrumental homicides, and especially mafia murders, concentrate in Southern Italy (Aziani, 2022). Mafia murders are frequently committed by a group, and the identity of the killers is often unknown to the mafia members themselves. Moreover, mafia groups can organize disinformation strategies to divert the police investigation, and by means of intimidation and threats they may induce witnesses not to cooperate with the police (Paoli, 2003). Hence the presence of structured criminal organizations may impact on homicide clearance. At the same time, the fact that the correlation between case clearance and macro-area remains significant despite the simultaneous control for the motivation behind the homicide indicates that other regional specificities play a role. Whether differences in investigation techniques and organizational strategies influence these discrepancies in clearance rates between Southern Italy and the rest of the country might be focus of future studies.

Post-estimation tests and analyses of the residuals signaled that the logistic regression models proposed did not violate the assumption behind their use (see Annex). The operationalization of regressors as previously done in the literature on clearance further advocates in favor of our models. Nonetheless, the proposed analysis has limitations that should be considered in interpreting the results. Due to information availability, we performed no systematic analysis of the role played by investigative activities, although investigative factors are known to be important in homicide clearance (see Braga &

Dusseault, 2018). Secondly, the classification through media analysis of certain variables is not indisputable. We mitigated this issue by relying also on a survey conducted on investigators who worked on the cases, and by cross-validating data gathered through media analysis. The aggregation into macro classes of factors, such as age and socioeconomic status, makes the evaluation of their association with homicides clearance less sensitive; however, these aggregations are necessary to limit the number of classes used in multivariate analyzes. With respect to machine learning exploratory analyses, the proposed inductive approach cannot provide evidence of causal inference. Therefore, we cannot assert that individual domain factors are more causally related to clearance than non-discretionary factors, we can only suggest that they are more important for predicting homicide clearance. Finally, the findings are not fully generalizable because the study focuses on Italy.

Conclusion

This study has assessed the extent to which victim's individual factors related to police discretion perspectives on homicide clearance and the characteristics and the circumstances of the criminal event may affect the outcome of a homicide investigation. The empirical analysis was conducted by applying Pearson's χ^2 test, logistic regressions, and random forest methods to Italian data. The evidence does not support the police's discretionary perspective. The proposed analyses did not identify victim's nationality, previous convictions, and socioeconomic status as significant predictors of the investigation's success. On the contrary, victim's sex is associated with homicide clearance. Moreover, media attention positively correlates with homicide clearance. Homicides occurring within a criminal environment were significantly less likely to be cleared, although the victim's criminal history does not impact investigators' chances to identify the murderer. Finally, the clearance rate was lower in Southern Italy; the ability of mafia groups to mislead investigators is likely to contribute to this result.

References

- Addington, L. A. (2006). Using National Incident-Based Reporting System Murder Data to Evaluate Clearance Predictors: A Research Note. *Homicide Studies*, 10(2), 140–152.
<https://doi.org/10.1177/1088767905285439>
- Addington, L. A. (2007). Hot vs. Cold Cases: Examining Time to Clearance for Homicides Using NIBRS Data. *Justice Research and Policy*, 9(2), 87–112. <https://doi.org/10.3818/JRP.9.2.2007.87>
- Alderden, M. A., & Lavery, T. A. (2007). Predicting Homicide Clearances in Chicago: Investigating Disparities in Predictors Across Different Types of Homicide. *Homicide Studies*, 11(2), 115–132.
<https://doi.org/10.1177/1088767907300505>
- Amin, A. (2013). Land of strangers. *Identities*, 20(1), 1–8. <https://doi.org/10.1080/1070289X.2012.732544>
- Aziani, A. (2022). *Violenza e mercati illegali: Teorie e analisi della realtà italiana*. Vita e Pensiero.
- Bänziger, M., & Killias, M. (2014). Unsolved homicides in Switzerland: Patterns and explanations. *European Journal of Criminology*, 11(5), 619–634. <https://doi.org/10.1177/1477370814541338>
- Black, D. J. (1976). *The Behavior of Law*. Academic Press.
- Braga, A. A., & Dusseault, D. (2018). Can Homicide Detectives Improve Homicide Clearance Rates? *Crime & Delinquency*, 64(3), 283–315. <https://doi.org/10.1177/0011128716679164>
- Braga, A. A., Turchan, B., & Barao, L. (2019). The Influence of Investigative Resources on Homicide Clearances. *Journal of Quantitative Criminology*, 35(2), 337–364. <https://doi.org/10.1007/s10940-018-9386-9>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32.
<https://doi.org/10.1023/A:1010933404324>
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Wadsworth & Brooks/Cole Advanced Books & Software.

- Aziani and Persurich (2023) Homicide clearance: Discretionary and non-discretionary factors (preprint version). Publication available at: <https://journals.sagepub.com/doi/epub/10.1177/14773708221136049> - alberto.aziani@unimib.it – alberto.aziani@unicatt.it
- Brookman, F., Maguire, E. R., & Maguire, M. (2019). What Factors Influence Whether Homicide Cases Are Solved? Insights From Qualitative Research With Detectives in Great Britain and the United States. *Homicide Studies*, 23(2), 145–174. <https://doi.org/10.1177/1088767918793678>
- Campedelli, G. M. (2022). Explainable machine learning for predicting homicide clearance in the United States. *Journal of Criminal Justice*, 79, 101898. <https://doi.org/10.1016/j.jcrimjus.2022.101898>
- Carter, D. L., & Carter, J. G. (2016). Effective Police Homicide Investigations: Evidence From Seven Cities With High Clearance Rates. *Homicide Studies*, 20(2), 150–176. <https://doi.org/10.1177/1088767915576996>
- Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton University Press.
<https://press.princeton.edu/books/paperback/9780691005478/mathematical-methods-of-statistics-pms-9-volume-9>
- Gottfredson, M. R., & Hindelang, M. J. (1979). A Study of the Behavior of Law. *American Sociological Review*, 44(1), 3–18. <https://doi.org/10.2307/2094813>
- Graversen, T., Mortera, J., & Lago, G. (2019). The Yara Gambirasio case: Combining evidence in a complex DNA mixture case. *Forensic Science International: Genetics*, 40, 52–63.
<https://doi.org/10.1016/j.fsigen.2018.12.010>
- Hawk, S. R., & Dabney, D. A. (2014). Are All Cases Treated Equal?: Using Goffman’s Frame Analysis to Understand How Homicide Detectives Orient to Their Work. *The British Journal of Criminology*, 54(6), 1129–1147. <https://doi.org/10.1093/bjc/azu056>
- Hosmer, D. W. Jr., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression*. John Wiley & Sons.
- Hothorn, T., Seibold, H., & Zeileis, A. (2021). *partykit: A Toolkit for Recursive Partytioning* (1.2-15).
<https://CRAN.R-project.org/package=partykit>
- Hothorn, T., & Zeileis, A. (2015). partykit: A Modular Toolkit for Recursive Partytioning in R. *Journal of Machine Learning Research*, 16(118), 3905–3909.

- Aziani and Persurich (2023) Homicide clearance: Discretionary and non-discretionary factors (preprint version). Publication available at: <https://journals.sagepub.com/doi/epub/10.1177/14773708221136049> - alberto.aziani@unimib.it – alberto.aziani@unicatt.it
- Hough, R., McCorkle, K. D., & Harper, S. (2019). An Examination of Investigative Practices of Homicide Units in Florida. *Homicide Studies*, 23(2), 175–194. <https://doi.org/10.1177/1088767919828421>
- Jiao, A. Y. (2007). Explaining Homicide Clearance: An Analysis of Chicago Homicide Data 1965–1995. *Criminal Justice Studies*, 20(1), 3–14. <https://doi.org/10.1080/14786010701241259>
- Keel, T. G., Jarvis, J. P., & Muirhead, Y. E. (2009). An Exploratory Analysis of Factors Affecting Homicide Investigations: Examining the Dynamics of Murder Clearance Rates. *Homicide Studies*. <https://doi.org/10.1177/1088767908326903>
- Klinger, D. A. (1997). Negotiating Order in Patrol Work: An Ecological Theory of Police Response to Deviance. *Criminology*, 35(2), 277–306. <https://doi.org/10.1111/j.1745-9125.1997.tb00877.x>
- LaFree, G., Curtis, K., & McDowall, D. (2015). How effective are our ‘better angels’? Assessing country-level declines in homicide since 1950. *European Journal of Criminology*, 12(4), 482–504. <https://doi.org/10.1177/1477370815584261>
- Lee, C. (2005). The value of life in death: Multiple regression and event history analyses of homicide clearance in Los Angeles County. *Journal of Criminal Justice*, 33(6), 527–534. <https://doi.org/10.1016/j.jcrimjus.2005.08.002>
- Liem, M., Krüsselmann, K., & Eisner, M. (2020). From Murder to Imprisonment: Mapping the Flow of Homicide Cases—A Systematic Review. *Homicide Studies*, 24(3), 220–241. <https://doi.org/10.1177/1088767920924447>
- Liem, M., Suonpää, K., Lehti, M., Kivivuori, J., Granath, S., Walser, S., & Killias, M. (2019). Homicide clearance in Western Europe. *European Journal of Criminology*, 16(1), 81–101. <https://doi.org/10.1177/1477370818764840>
- Litwin, K. J. (2004). A Multilevel Multivariate Analysis of Factors Affecting Homicide Clearances. *Journal of Research in Crime and Delinquency*, 41(4), 327–351. <https://doi.org/10.1177/0022427803260269>

- Aziani and Persurich (2023) Homicide clearance: Discretionary and non-discretionary factors (preprint version). Publication available at: <https://journals.sagepub.com/doi/epub/10.1177/14773708221136049> - alberto.aziani@unimib.it – alberto.aziani@unicatt.it
- Litwin, K. J., & Xu, Y. (2007). The Dynamic Nature of Homicide Clearances: A Multilevel Model Comparison of Three Time Periods. *Homicide Studies*, 11(2), 94–114. <https://doi.org/10.1177/1088767907300759>
- Mancik, A. M., Parker, K. F., & Williams, K. R. (2018). Neighborhood Context and Homicide Clearance: Estimating the Effects of Collective Efficacy. *Homicide Studies*, 22(2), 188–213. <https://doi.org/10.1177/1088767918755419>
- Mouzos, J., & Muller, D. (2001). *Solvability factors of homicide in Australia: An exploratory analysis*. Australian Institute of Criminology.
- Nagin, D. S. (1998). Criminal Deterrence Research at the Outset of the Twenty-First Century. *Crime and Justice*, 23, 1–42. <https://doi.org/10.1086/449268>
- Ousey, G. C., & Lee, M. R. (2010). To Know the Unknown: The Decline in Homicide Clearance Rates, 1980–2000. *Criminal Justice Review*, 35(2), 141–158. <https://doi.org/10.1177/0734016809348360>
- Paoli, L. (2003). *Mafia Brotherhoods: Organized crime, Italian Style*. Oxford University Press.
- Paternoster, R. (1984). Prosecutorial Discretion in Requesting the Death Penalty: A Case of Victim-Based Racial Discrimination. *Law & Society Review*, 18(3), 437–478. JSTOR. <https://doi.org/10.2307/3053431>
- Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302), 157–175. <https://doi.org/10.1080/14786440009463897>
- Petersen, N. (2017). Examining the Sources of Racial Bias in Potentially Capital Cases: A Case Study of Police and Prosecutorial Discretion. *Race and Justice*, 7(1), 7–34. <https://doi.org/10.1177/2153368716645842>
- Peterson, R. D., & Hagan, J. (1984). Changing Conceptions of Race: Towards an Account of Anomalous Findings of Sentencing Research. *American Sociological Review*, 49(1), 56–70. <https://doi.org/10.2307/2095557>

- Aziani and Persurich (2023) Homicide clearance: Discretionary and non-discretionary factors (preprint version). Publication available at: <https://journals.sagepub.com/doi/epub/10.1177/14773708221136049> - alberto.aziani@unimib.it – alberto.aziani@unicatt.it
- Pizarro, J. M., Terrill, W., & LoFaso, C. A. (2020). The Impact of Investigation Strategies and Tactics on Homicide Clearance. *Homicide Studies*, 24(1), 3–24. <https://doi.org/10.1177/1088767918816741>
- Puckett, J. L., & Lundman, R. J. (2003). Factors Affecting Homicide Clearances: Multivariate Analysis Of A More Complete Conceptual Framework. *Journal of Research in Crime and Delinquency*, 40(2), 171–193. <https://doi.org/10.1177/0022427803251125>
- Regoeczi, W. C., Jarvis, J., & Mancik, A. (2020). Homicide Investigations in Context: Exploring Explanations for the Divergent Impacts of Victim Race, Gender, Elderly Victims, and Firearms on Homicide Clearances. *Homicide Studies*, 24(1), 25–44. <https://doi.org/10.1177/1088767918802885>
- Regoeczi, W. C., Kennedy, L. W., & Silverman, R. A. (2000). Uncleared Homicides: A Canada/United States Comparison. *Homicide Studies*, 4(2), 135–161. <https://doi.org/10.1177/1088767900004002002>
- Riedel, M. (2008). Homicide Arrest Clearances: A Review of the Literature. *Sociology Compass*, 2(4), 1145–1164. <https://doi.org/10.1111/j.1751-9020.2008.00130.x>
- Riedel, M., & Jarvis, J. (1999). The Decline of Arrest Clearances for Criminal Homicide: Causes, Correlates, and Third Parties. *Criminal Justice Policy Review*, 9(3–4), 279–306. <https://doi.org/10.1177/088740349900900302>
- Ripley, B. D. (1996). *Pattern Recognition and Neural Networks*. Cambridge University Press.
- Roberts, A. (2007). Predictors of Homicide Clearance by Arrest: An Event History Analysis of NIBRS Incidents. *Homicide Studies*, 11(2), 82–93. <https://doi.org/10.1177/1088767907300748>
- Roberts, A., & Lyons, C. J. (2009). Victim-Offender Racial Dyads and Clearance of Lethal and Nonlethal Assault. *Journal of Research in Crime and Delinquency*, 46(3), 301–326. <https://doi.org/10.1177/0022427809335168>
- Roberts, A., & Lyons, C. J. (2011). Hispanic Victims and Homicide Clearance by Arrest. *Homicide Studies*, 15(1), 48–73. <https://doi.org/10.1177/1088767910397278>

Aziani and Persurich (2023) Homicide clearance: Discretionary and non-discretionary factors (preprint version). Publication available at: <https://journals.sagepub.com/doi/epub/10.1177/14773708221136049> - alberto.aziani@unimib.it – alberto.aziani@unicatt.it

Rydberg, J., & Pizarro, J. M. (2014). Victim Lifestyle as a Correlate of Homicide Clearance. *Homicide Studies*, 18(4), 342–362. <https://doi.org/10.1177/1088767914521813>

Strobl, C., Boulesteix, A.-L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, 8(1), 25. <https://doi.org/10.1186/1471-2105-8-25>

Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods*, 14(4), 323–348. <https://doi.org/10.1037/a0016973>

Vaughn, P. E. (2020). The effects of devaluation and solvability on crime clearance. *Journal of Criminal Justice*, 68, 101657. <https://doi.org/10.1016/j.jcrimjus.2020.101657>

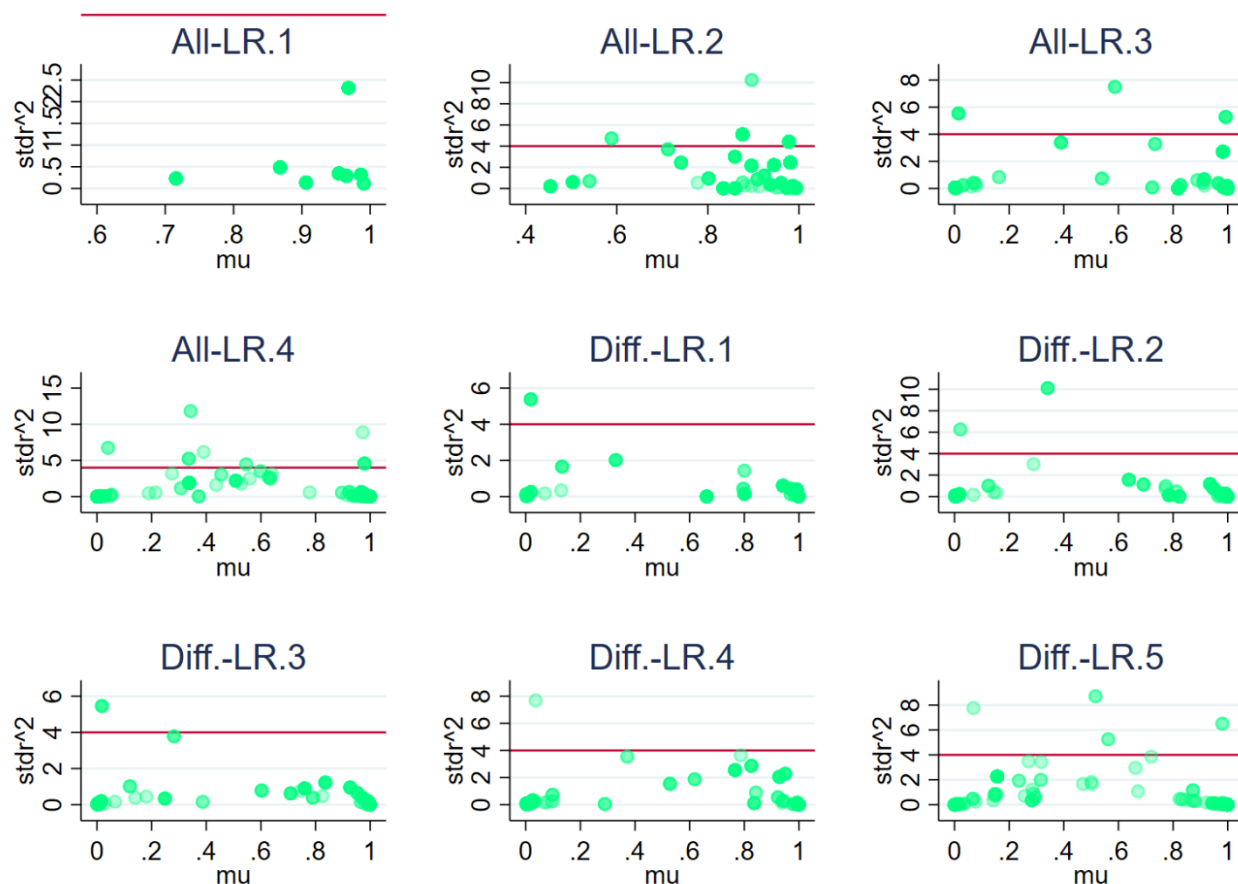
Wellford, C., & Cronin, J. (1999). *Analysis of Variables Affecting the Clearance of Homicides: A Multisite Study* (No. 181356; p. 175). Justice Research and Statistics Association. <https://www.ojp.gov/library/publications/analysis-variables-affecting-clearance-homicides-multisite-study>

Zanfrini, L. (2019). *The Challenge of Migration in a Janus-Faced Europe*. Palgrave Pivot. <https://doi.org/10.1007/978-3-030-01102-4>

Annex: Post-estimation analyses

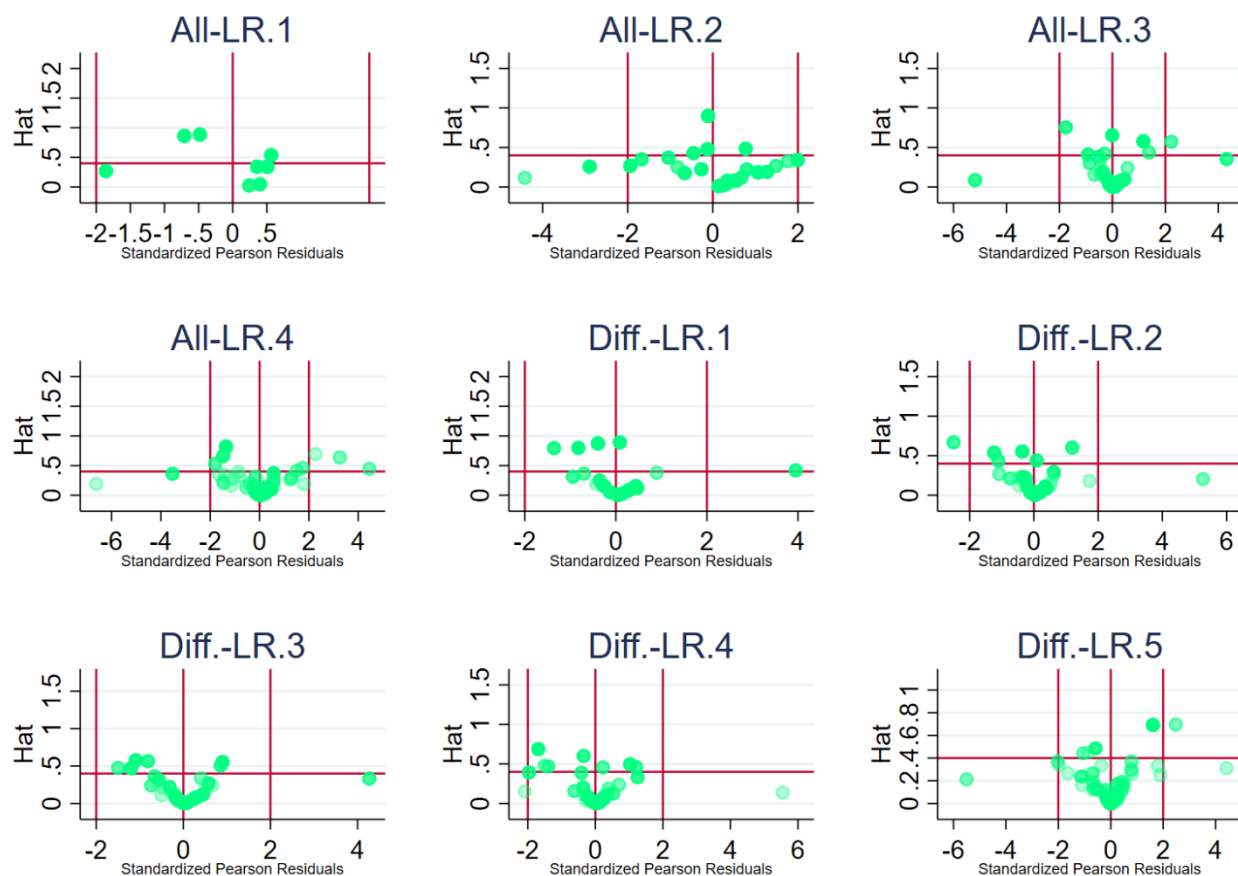
The analysis of residuals plays an important role in evaluating the possible influence of outliers and in assessing logistic model fit. First, we inspected the graphs of the squared standardized deviance versus μ (Figure A1). This inspection indicated that the analysis was not severely affected by the presence of outliers. Secondly, we investigated large *hat* values indicating covariate patterns that differed from average covariate patterns (Figure A2). Values that do not fit the model well can be identified by selecting *hat* values greater than 0.4 and squared residual values of $|2|$. In model All-LR.4 Difficult-LR.2 Difficult-LR.5, some observations fitted this characterization; but, overall, also this analysis confirmed the goodness of fit of the proposed models.

Figure A1. Squared Standardized Deviance versus predicted probability of clearance



Note: $stdr^2$ stands for Squared Standardized Deviance of Residuals; μ represents the predicted probability of a homicide being cleared. Values greater than 4 in the $stdr^2$ axis can be considered as outliers.

Figure A2. Squared Standardized Deviance versus predicted probability of clearance



Note: *hat* denotes diagonal elements of the projection matrix. Large *hat* values indicate covariate patterns that differ from average covariate patterns. Observations characterized by *hat* values greater than 0.4 and standardized Pearson residuals below -2 or above +2 do not fit the model well.