Crime is Terribly Revealing: Information Technology and Police Productivity*

Giovanni Mastrobuoni†

December 2014‡

Abstract

In an unprecedented information technology (IT) revolution in the public service sector, an increasing number of police departments use advanced statistical methods to improve their productivity in fighting crimes. Since 2007 the police department of Milan has been using a predictive policing software that is unique, as it not only produces aggregate crime predictions but also individual ones.

This paper uses detailed information on *individual* crime incidents, coupled with offender-level identifiers produced by the software, to show that criminals follow habits, and that such habits make their future actions predictable. Using quasi-random assignment of crimes to two police forces that differ in the availability of this predictive policing software, this study shows that the adoption of this very advanced yet inexpensive IT innovation doubles the productivity of policing.

Keywords: predictive policing, police, crime, quasi-experiment
JEL classification codes: O33, K42, L23, H1, H41

*I would like to thank the Police Chief of Milan (*Questore di Milano*) for providing the data, as well as Mario Venturi and his staff for sharing their knowledge on robberies and policing with me. I would also like to thank Mark Kleinman, Emily Owens, David Rivers, and Arie Kapteyn as well as seminar participants at the Collegio Carlo Alberto, the 2013 Al Capone workshop, and the 2013 NBER Summer Institute for invaluable comments.

†Department of Economics, University of Essex, Collegio Carlo Alberto and Netspar, Email: gmastrob@essex.ac.uk.

"But, yes, Hastings, I think it is almost certain there will be another. A lot depends on *la chance*. So far our *inconnu* has been lucky. This time the luck may turn against him. But in any case, after another crime, we shall know infinitely more. **Crime is terribly revealing.** Try and vary your methods as you will, your tastes, your habits, your attitude of mind, and your soul is revealed by your actions. There are confusing indications - sometimes it is as though there were two intelligences at work - but soon the outline will clear itself, I shall know." (Agatha Christie, 1936)

1 Introduction

Over the past 30 years, service organizations, including those in the public sector, have shown a dramatic increase in the use of information technologies (IT). While a large body of research investigated the relationships between IT, work organization, and productivity,[1] only a few studies show direct evidence about the role of IT in increasing service sector productivity, including in the public sphere.[2]

IT investments are often intangible and disproportionately difficult to measure (see Brynjolfsson and Hitt, 2000, David, 1990), and even when detailed data are available, estimates of IT impacts are usually based on cross-sectional or at best panel-data variation in IT use. But organizations that use IT innovations may be those that benefit the most from such innovations or differ in ways that are unobserved to the econometrician. Even when focusing on IT adoption in a very narrowly defined industry (see Bartel et al., 2007), such adoption might coincide with other new management practices (Athey and Stern, 2002). Finally, even if all practices are measured with care, their adoption, including IT,

[1]See, among others, Acemoglu et al. (2007), Autor et al. (1998), Berman et al. (1994), Black and Lynch (2001), Bresnahan et al. (2002), Doms et al. (1997), Stiroh (2002).

[2]See Angrist and Lavy (2002), Athey and Stern (2002), Goolsbee and Guryan (2006), and Garicano and Heaton (2010).

might still be driven by unobserved factors. This study overcomes these issues by focussing on a very specific IT innovation (predictive policing), that as I argue later, appears to be as good as randomly assigned to one of two separate public service providers (police departments) that operate side by side in the same city.

I examine the relationship between the use of advanced IT strategies and the productivity of policing, measured by the likelihood that a crime is cleared (solved) by arrest.[3]

Most US police departments employ geographic information systems (GIS) to map and analyze crime trends across different geographic areas, and use such information to optimize their patrolling strategies (Garicano and Heaton, 2010, Weisburd et al., 2003).[4] Such practices are often summarized by "Compstat," which was developed by the New York Police Department's Police Commissioner William Bratton under Mayor Rudolph Giuliani's leadership.

The reason for such practices resides in a striking empirical regularity: few intersections, or city blocks often produce the majority of crime incidents, called crime hot spots (see, among others, Sherman et al., 1989, Weisburd and Eck, 2004, Weisburd and Green, 1995). These patterns have prompted police departments to target policing in geographic areas (e.g., blocks or specific addresses) that show high levels of criminal activity. Several studies have evaluated hot-spots policing strategies, and, at least among criminologists, the general consensus is that focussed policing works (Braga, 2001, Cohen and Ludwig, 2003, Sherman and Weisburd, 1995, Weisburd and Green, 1995).[5]

There is also currently little evidence that hot-spots policing simply displaces crime to

[3]Several economic studies have used clearances as a measure of police performance (see, among others, Garicano and Heaton, 2010, Mas, 2006). This paper is broadly related to the literature that estimates the effect of police on crime (see, among others Buonanno and Mastrobuoni, 2011, Corman and Mocan, 2000, Evans and Owens, 2007, Levitt, 1997, Machin and Marie, 2011). There is increasing evidence that increasing police investments reduces crime, but little is known about the underlying production process (Cook et al., 2011, Levitt and Miles, 2004).

[4]Garicano and Heaton (2010) studies the relationship between information technology, productivity, and the organization of police departments. Such investments are linked to improved productivity when they are complemented with programs like Compstat.

[5]Levitt (2004) is more skeptical about the decline in crime that occurred during the 1990s that can be attributed to Compstat.

nearby locations (some studies even find a reduction in crime in the surrounding locations,[6] though one potential limitation of these studies is identifying the area where crime might spill over, for this area is not necessarily contiguous to the area that is being targeted (McCrary, 2010).

Hot-spots policing has gradually evolved from using data to simply identify high crime areas into a more dynamic and advanced information technology that uses these data to make predictions about future aggregate criminal activity. Past aggregate crime patterns have been shown to predict the most likely type, location, and time of future crimes (Mohler et al., 2011).

Despite growing interest in predictive policing, very little is known about its effectiveness, and why it might work. Indeed, an open question is whether predictive policing is capable of improving the performance of policing. As for hot-spots policing, the main burdens are: i) the non-random assignment of the use of predictive policing to crimes; ii) the additional deterrent effect that might in principle displace crime across time and space, potentially overstating the benefits of predictive policing; and iii) the additional incapacitation effect might in principle spill over to other regions, potentially understating the benefits of predictive policing. In Section 4 it will be shown that aggregate crime data and aggregate predictions cannot overcome these burdens.

In order to better understand why predictive policing (as well as proactive policing more generally) might work, and to reveal the nature of the spillovers, one needs to uncover the mechanism that makes crimes predictable across time and space. Such predictability might be driven by the characteristics and the activity of criminals (e.g. habits, place of residence, commuting patterns, etc.), the characteristics and activity of victims (e.g. victims' location, place of residence, hours of "exposure," etc.), and the interplay between criminals and victims (e.g. gang shootings, victim's precautions, etc.).[7]

[6] See Clarke and Weisburd (1994) and Weisburd et al. (2006).
[7] See Clarke (1997) for a discussion about situations that enhance criminal opportunities.

This study uses a unique *micro-level data* on commercial robberies, which are crimes of violence against businesses motivated by theft, and predictions about *individual* crimes to uncover the mechanism underlying predictive policing.[8][9].

The Milan Police Department started collecting and analyzing these data in 2007, when a number-crunching police officer noticed that the robberies that could be linked to repeat offenders had some common attributes. Realizing that data collection and analysis could potentially lead to a more accurate policing, in his spare time he developed and later copyrighted the software "KeyCrime." "KeyCrime" analyzes large sets of individual characteristics of robbers and individual *modus operandi* in order to i) identify robberies that share at least one offender (called a "series"); and ii) predict when and where the "series" is going to continue. The software has been leased at no cost to the Milan Police Department. The linkages across robberies are constructed irrespective of whether an arrest is made, because they are based on the characteristics of the robbers (see Section 2.2).[10]

The police force uses this information for two purposes. The first is to predict the time and location of future commercial robberies, thereby optimizing police patrolling. The second is to assist the prosecutors once the perpetrators have been arrested and are put on trial.[11]

[8]According to the US Uniform Crime Reports in 2009 robbery rates were 133 per 100,000 inhabitants, while they were 58.7 per 100,000 inhabitants in Italy (Barbagli and Colombo, 2011). About 25 and 42 percent of robberies reported to the police occur in businesses in Italy and in the US (Barbagli and Colombo, 2011, Cook, 2009).

[9]Mastrobuoni and Owens (2014) show that robbery rates in Milan are similar to the ones in other Italian cities, and that robbery rates in Italy are similar in magnitude to what happens in the US, Canada, and the UK.

[10]A first thing to notice is that such bottom-up innovations are arguably less prone than top-up ones to be an endogenous response to increasing crime rates. Indeed, the Milan Police Department is the only one out of more than 200 departments across 102 provinces to have developed such a predictive software. While the assignment of predictive policing seems to be purely accidental, the identification strategy that I develop does not rest on this assumption.

[11]Thus not only clearance rates, defined as the likelihood of solving a specific crime before the offender's next crime, are likely to respond to this IT innovation; conviction rates could potentially improve as well. Unfortunately, the identification strategy used to estimate the causal effect of predictive policing on clearance rates cannot be extended to conviction rates. The reason is that all police forces share all information collected with the prosecutors, even when the competing police force, the *Carabinieri*, which

I set up a difference-in-difference identification strategy that exploits i) the quasi-random assignment of investigations of robberies against businesses to two separate police forces (*Polizia* or police and *Carabinieri* or gendarmerie) due to a very peculiar rotating mechanism that four times a day (during shift changes) forces the two police forces to cover different parts of the city; ii) the very nature of predictive policing, which is based on the analysis of past crimes.

I show that predictive policing doubles the likelihood of clearing a robbery (by around 8 percentage points), but not for the very first robbery of a series, for which no past data are available.[12] The identification rests on the assumption that differences in clearance rates between the two police forces that are not driven by predictive policing are the same for first and subsequent robberies.[13] Errors in linkages, which are likely, lower the accuracy of the predictions and the efficacy of predictive policing, but do not bias the estimates.

A second identification strategy exploits the time it takes the police to collect and analyze the data. Predictions are never updated the same day that a new robbery takes place. Computing difference-in-difference in clearance rates between the two police forces for "same day" linked robberies and "subsequent day" linked robberies shows even larger treatment effects, though I will argue that this estimate represents an upper bound of the IT impact on police performance. The gendarmerie shows no differences in clearance rates between same day and subsequent day robberies, which is consistent with differences being driven by the delayed predictions.

Given that both identification strategies are based on contemporaneous comparisons

---

later represents our control group, made the arrest.

[12]With the exception of predictive policing, the two police forces share similar staffing and equipment (see Section 4.1). This is likely why the two forces have the same likelihood of clearing first robberies (around 12 percent).

[13]In principle, when analyzing photographic evidence the police might recognize individuals even if they have never been arrested before. Since later for first robberies I do not find significant differences in clearance rates between the *Polizia* and the *Carabinieri*, either this happens very rarely, or both forces look at the photographic evidence.

between two forces, one that innovated and one that did not, this paper highlights short term benefits based on improved productivity of policing. On top of the incapacitation effect that is driven by arrest and convictions (all arrests but one lead to convictions), as criminals learn about the improved productivity of the police, a deterrence effect might push criminals towards other crimes, or other cities. The identification strategy employed in this paper cannot be used to estimate this deterrence effect, and so I leave this task for future research. Moreover, crime reductions may be due to deterrence and/or incapacitation.[14] Under certain assumptions the impact on clearances translates into more incapacitation, which I quantify. A simple cost-benefit analysis indicates that predictive policing can generate large societal benefits (see Section 6).

Apart from showing that predictive policing works, this study provides some preliminary evidence of why it works. While I was not given access to the algorithm that predicts criminal behavior, I show that individual criminal behavior is in part predictable.[15] Over time criminal groups tend to select the same business types, around the same time of the day, and in the same city neighborhood, especially if previous robberies have been lucrative. I provide evidence that the instructions distributed to the police patrols take these patterns into account.

[14]The two mechanisms are often hard to separate when only aggregate data are available (Owens, 2014). See Durlauf et al. (2010) for additional issues that might arise from estimating aggregated crime regressions. Mastrobuoni uses the same crime level data used in this paper, in particular the variation in police presence that is driven by shift changes, to show that an increase in police patrolling leads to higher clearance rates. Di Tella and Schargrodsky (2004), Draca et al. (2011) and Klick and Tabarrok (2005) exploit exogenous variation is the deployment of "high deterrence" police officers following terrorist attacks, and find strong evidence in favor of a deterrent effect of police stationing a circumscribed area.

[15]Predictability does not necessarily mean that criminals are not choosing an optimal criminal strategy. Becoming more unpredictable seems costly: apart from the potential cost of travelling more, the data shows that targeting different types of businesses is associated with a lower loot.

2 Predictive Policing

2.1 Around the World

The precursor of predictive policing is Compstat, a data gathering and accountability process started by the New York Police Department in 1995 and since then adopted by most US police departments (Weisburd et al., 2003).

Predictive policing is also based on GIS data, but uses more advanced statistical techniques that are built on autoregressive models over time and space. The most advanced ones predict the most likely type, location, and time of future crimes.[16]

Recently the National Institute of Justice (NIJ) has launched a demonstration initiative to develop, test and evaluate predictive policing in a real-world, real-time context and awarded planning grants to several law enforcement agencies (Pearsall, 2010).[17]

The Chicago Police Department is partnering with computer scientists at the Illinois Institute of Technology to develop a crime-fighting algorithm. In Memphis, IBM is part of a project called Blue CRUSH (Criminal Reduction Utilizing Statistical History). But only in 2011 did the first US department evaluate predictive policing. The Santa Cruz police department ran a city-wide 6 months "Predictive Policing Experiment," named one of Time Magazine's 50 best innovations of 2011 (Grossman et al., 2011). Like many police departments around the world, the Santa Cruz Police Department had a declining budget and shrinking police force. After an unprecedented crime wave at the beginning of 2011 the department decided to work with researchers at UCLA to test a new method of modelling crime using data on 2,803 burglaries (Economist, 2010, Mohler et al., 2011).[18] The experiment seemed to reduce crime, though the absence of a control group and

[16] The PredPol company uses "self-exciting point process modeling," where decreasing kernels are used to weight the observations that are farther away in space and time (Mohler et al., 2011).

[17] The list of seven police departments is: Los Angeles, Boston, Chicago, Maryland State, Richmond, Las Vegas, District of Columbia Metropolitan and Shreveport. Two of the original seven sites (Chicago and Shreveport) won competitively awarded grants to continue into Phase 2 of their demonstration and evaluation of predictive policing strategies.

[18] The experiment is described here http://math.scu.edu/~gmohler/predpol.html.

the possibility that crime was merely displaced make it difficult to draw some definite conclusions.[19]

The issue in most of these studies is that they either lack a control region, or that criminals might be moving from treated to control regions contaminating the experimental design. Reducing contamination by choosing larger regions and exploiting pure time-series variation would also be unpractical. A spike in crime followed by the use of predictive policing might, just naturally, lead to reversion to the mean that is completely unrelated to the newly adopted technology. Moreover, part of the effect of predictive policing might be due to an incapacitation effect, which is dynamic in nature, and thus hard to separate over time.

Later in Section 4, I will argue that if criminals are mobile, one cannot use a fixed treatment assignment, in either time or space, to evaluate predictive policing; especially, if criminals perceive changes in the productivity of the police forces. The same argument applies to other policing strategies, like hot spot policing.

The predictive policing adopted by the Milan Police Department and the Italian institutional background overcome these issues.

2.2 The Milan Police Predictive Policing

The IT innovation used by the Milan police collects and later examines around 11,000 bits of information about each robbery (time, date, location, type of business, type of crime, etc.), about the observed criminals (perceived age, height, body structure, skin, hair, eye color, clothing, etc.), about the observed weapons (type, maker, model, color, etc.), and about the observed vehicle used (type, maker, model, license-plate, etc.).[20] In particular,

[19]Predictive policing is also being evaluated in the UK where, in the single ward of the Greater Manchester area studied, burglary decreased by 26 percent versus 9 percent city-wide, which led to follow-up studies in Birmingham.

[20]According to the police no other major innovation has been adopted by either the gendarmerie or the police. However, later in the analysis I am going to allow for differential productivity (including innovations) between the police and the gendarmerie, as long as such differences in productivity apply to

after each reported robbery–by June 2011 the Milan Police department had recorded around 2000 robberies, at a rate of 1.5 robberies per day–the police force gathers data about the event collecting the official police reports, interviewing victims and collecting surveillance camera footage.[21]

Driven by incentives to collect insurance money and to increase future protection from police patrols, reporting rates among commercial businesses are basically equal to one.[22]

The victim interviews, which represent the core of the information collected, happen mostly over the phone at least 24 hours from the time of the crime to reduce the victim's post-traumatic stress disorder.[23] These data are used to establish links across robberies (see Figure 1), highlight regularities, and use these regularities to predict future potential targets. Even in the absence of photographic evidence, the mutual appearance of one or several peculiar elements that characterize a robbery or a robber might help establishing a link.

The victimized targets and the potential future targets (some are indicated in Figure 3 with a small blue square) are later communicated to police patrols, together with the likely day of the week and time of the day of the future offense. An example of such a report is shown in Figure 2. The report describes the offenders and their typical *modus operandi*, including the means of transportation, the typical time of the day and target type chosen. On the second page of the report a map indicates the neighborhoods where the criminals are likely to strike, while the final page collects all the photographic evidence. The group of criminals shown in Fig. 2 has presumably robbed 22 business, which is why such evidence is particularly rich.

Such a micro-based predictive approach is quite unique. For this reason, while po-

all robberies.

[21]While not all businesses have closed-circuit security cameras, most banks, postal offices, and jewelers typically do.

[22]According to the police, among thousands of robberies, only in one instance did a robber confess to a robbery that had not been reported before. The business owner confirmed that he had not reported the crime.

[23]Later I exploit such delays to setup one difference in differences strategy.

tentially capable of providing new insights about how any predictive policing works, one should keep in mind the uniqueness of the approach developed in Milan when trying to generalize the evaluation of the Milan "micro-predictions" to more "macro" ones.

I have been given access to a subset of the data that "KeyCrime" processes, with great detail on the *modus operandi* of the robberies (location, time, loot, arrest, number of offenders, weapons, type of business, etc.). The Milan police also collects data on the physical characteristics of the offenders, as well as photographic footage, but these are not included in the data that was released to me. The summary statistics of the available variables are shown in Table 1, both for the full sample and for the sample which restricts the data to the first two years. Each observation represents a separate robbery. Over the period 2008-2011 there were over 2000 separate robberies in Milan. According to the Milan police 70 percent of these robberies show some link with other robberies, meaning that at least one robber, or one weapon, or one vehicle were seen in two different instances. The variable "Number of the series" $n = 1, ..., N_i$ measures the number of robberies each group of linked offenders i have been involved in.[24] The criminal group with the largest number of offences organized 84 robberies.

The police force defines a given robbery to be cleared if an arrest is made before the same group of robbers re-offends.[25] More than half of the robberies (1,221 robberies out of 2,164) belong to a series where at least one arrest has been made. Of these, 981 (80 percent) belong to a series that has been presumably fully cleared.[26]

An uncleared robbery ($C_{i,n} = 0$), within a series of robberies where some robberies have been cleared ($C_{i,k} = 1, k \neq n$), signals that either some perpetrators have not yet been identified or that there was insufficient evidence to attribute the robbery to some

[24]For the series that started in 2007 and continued in 2008 I was given the number of robberies performed in 2007, which I added to the "Number of the series."

[25]I do not have complete information on the exact date of arrest, but according to the police the majority of arrests happen *in flagrante*, meaning just before (when the police recognizes the criminals), during, or just after the robbery has taken place.

[26]Though, it would still be possible for the series to proceed if new perpetrators were using the same weapon or the same vehicles used by the arrested ones.

of the robbers that participated to the robberies. I define a series i to be solved when all robberies $n = 1, ..., N_i$ appear to be cleared. Finally, I define a single robbery to be "newly" cleared ($C_{i,n} = 1$) whenever the last robbery has been cleared ($C_{i,N_i} = 1$, there are 279 such robberies) or when a cleared robbery is preceded by an uncleared one ($C_{i,n-1} = 0, C_{i,n} = 1$, there are 21 such robberies).

Moreover, clearance patterns where an uncleared robbery is followed by two cleared ones (0,1,1) signals that a new arrest has been made in the second *and* in the third robbery.[27] For repeat offenders arrests happen most often just before a planned robbery, or immediately after an actual one. While the data cannot be used to reconstruct the arrests that happened just before a planned robbery, according to the police force more than half of all arrests happen immediately after the robbery takes place.

Table 1 shows that the individual clearance rate of robberies is 14.9 percent, which leads to 45 percent of the series being fully cleared by June 30, 2011, the day the data were extracted. The Police variable indicates whether the police handled that particular robbery and the next Section is going to describe how this assignment of investigations to the Police and the Gendarmerie works.

The robbers appear to be on average 26 years old. The average haul is around €2,000, or $2,600. One quarter of robberies are armed, and in about 10 percent of robberies a knife is used. Robberies are mainly an "Italian job," meaning that in 80 percent of cases at least one Italian seems to be involved. Only in 12 percent of cases the robbers seem to be of different nationalities. The average number of robbers involved in each robbery is about 1.5.

The next section describes what the ideal experiment to evaluate predictive policing would look like, and how the quasi-experiment resembles the ideal one.

[27]Given the rarity of 0-1-1 patterns the results are robust to defining the clearance rate based on just the last robbery.

3 Experimental Design

3.1 The Ideal Experiment

Given that many robbers are repeat offenders, it would be difficult to evaluate predictive policing, "the treatment," using differences over time. Some criminals, probably the more able ones, might learn about the availability of the new investigative techniques and as a response, switch to other crimes or move to other locations or points in time.

Moreover, part of the effect of predictive policing comes from incapacitation: preventing subsequent crimes by captured criminals. The presence of this dynamic reduction in crime would be more persistent in nature (lasting at least for the whole duration of the incarceration), making it hard to infer the effect of predictive policing from simple pre-post analyses. Using contemporaneous differences in treatment across locations would also be problematic. In order to see why, assume that location T has been treated and that location C represents the control region. Whenever criminals are mobile and target victims in both locations, an arrest in region T would also influence crimes in region C. Furthermore, more able criminals might perceive that in some areas policing is more productive, and select time and location of their offenses accordingly.

Given that the assignment of treatment across fixed regions is infeasible, the only other option is to assign treatments to changing regions but to the same patrols. Ideally one would exploit differences in treatment assignments within the same city and time, across different patrols. Some patrols would be given the instructions based on predictive policing software (like the ones shown in Fig. 2), and some would not. But the design would have to make sure that during the experimentation, police patrols are permanently assigned to either the treatment or to the control group, and that the officers in the two groups never interact. A "control" patrol might otherwise use some of the information collected when treated, or gather information by interacting with "treated" colleagues,

violating the so-called Stable Unit Treatment Value Assumption (SUTVA).

Finally, even with such a design, it would still be necessary to randomly assign crimes. If crimes were not randomly assigned to patrols an increased productivity (higher clearance rate) might be the sole product of selection. Otherwise, the increase in the likelihood of arrest might just be driven by treated patrols cherry-picking the more predictable and potentially less able criminals. If these types of criminals differed in their social dangerousness the evaluation might be misleading. For this reason an experiment would have to assign crimes to treatment and control patrols. The only way to do this is to assign areas to treatment and control patrols. But such an assignment would have to change over time and be unpredictable by criminals; otherwise one would go back to the issues about deterrence and spillovers discussed initially.

A final potential source of bias stems from the concept of the "Hawthorne effect," where an improvement in the performance might be driven by the mere attention given by the experimenter. In other words, the mere perception that one is participating in an experiment might generate a productivity response that is not related to the innovation per se. A possible solution to alleviate this concern would be to hide the existence of an experimental evaluation, have what is called a "blind" experiment.

Since one cannot track crime rates back to single police patrols (unless offenders concentrate their crimes in very short time periods and in small areas), a first corollary to the ideal experimental design we just described is that the evaluation needs to be based on the likelihood of clearing a crime that has been assigned to a given patrol rather than on crime rates.

A second corollary is that while the experiment might generate incapacitation and deterrence effects, the design would only identify the first. The unpredictability of treatment (necessary to avoid selection) would distribute the potential deterrence effects across both treatment and control patrols.

In summary, i) the experiment would have to be blind (Hawthorne effects); ii) treatment would have to be randomly (selection) and permanently assigned to patrols avoiding interaction between treatment and control groups (SUTVA); iii) the assignment of robberies to treatment and control patrols would have to be random (cherry picking); and iv) the assignment of robberies to treatment and control patrols would have to be unpredictable (endogenous response of criminals); and v) the productivity of the patrols would have to be measured in terms of their ability in solving a randomly assigned crime.

The Milan quasi-experiment is very close to such an ideal design.

4 The Quasi-experiment

4.1 Two Police Forces

For historical reasons, Italy has two separate police forces;[28] the *Carabinieri* is a military police force under the Italian ministry of defense and the *Polizia di Stato* is a civilian police force under the ministry of interior. The only difference between the two forces is that the police force operates exclusively in metropolitan areas, while the gendarmerie operates on the entire Italian territory. This difference is not going to influence this analysis as we are going to compare forces that operate within the boundaries of the city of Milan.[29]

The two forces share the same functions and objectives, which lead to considerable rivalry. Such rivalry leads to surprising commonalities. Not only do the two forces share the same equipment (e.g. the Beretta 92 is their standard service weapon, and the Alfa

[28]See Mastrobuoni for a discussion about the two forces.

[29]The *Carabinieri* might have an advantage when investigating criminal groups that operate both inside and outside of city. While the Milan police force does collect information on crimes that happen within the broader area of Milan (the province of Milan as opposed to the municipality) these crimes are always investigated by the gendarmerie. Moreover, the rotation mechanism that I am going to describe shortly, which gives rise to the quasi-random assignment, does not apply outside the metropolitan area (*Comune*). I do not have data on crimes that happen outside the city of Milan, but according to the *Polizia* the mobility of criminals in and out of the city should be quite limited.

Romeo 159, 2.4 JTDM 20v with 200 horsepower, is their standard service car, see Figure 4), they are almost identical in size. According to law, nationwide there are 57,336 police officers and 48,050 gendarmerie officers, both forces have 20,000 sergeants (*sovraintendenti*), they have almost the same number of inspectors (17,664 in the police and 16,031 in the gendarmerie), and the numbers of top-rank officials are similar as well.

But when investigating robberies they differ in the availability of predictive policing. I argued earlier that such availability is not an endogenous response to a crime wave or part of a wider set of innovations. It is rather a product of an enlightened number-crunching police officer, as well as, one should add, the Milan Police Department's decision to assist the police officer.

Up until the end of the 1990s the police and the gendarmerie were operating side by side, without communicating with each other.[30] Without a proper random assignment of crimes to police forces it would be impossible to estimate the productivity of the two police forces and the difference between the two, and relate such difference to differences in inputs. As I argued before, patrols that have access to predictive policing might target the more predictable criminals, inflating their productivity due to a selection of cases.

Fortunately for this study, at the end of the 1990s the government decided that to save resources the two forces would divide major Italian cities into different areas, and each force would be solely responsible for keeping law and order in a given area. In Milan the city is divided into three areas, two falling under the responsibility of the police and the third of the gendarmerie. Alone, even such division into areas would not provide random variation, because forces would probably be assigned to the zones according to their productivity, and criminals could potentially react by selecting the victims based on such assignment.

The additional variation I exploit is driven by a very peculiar rotation mechanism: the assignment of police forces to the three areas rotates approximately every 6 hours,

[30] The communication aspect started changing in January 2010.

counterclockwise (at 12am, 7am, 1pm, and 7pm). Given that there are two forces, three areas, and four 6-hour shifts within a given day, a given force is going to cover the same area during the same 6-hour shift only every three days. This means that there is quasi-random variation in the days of the month, days of the week, and 6-hour shift in the coverage of police forces. Figure 5 shows the distribution of robberies in Milan based on the day triplet, where the robberies that are under the responsibility of the gendarmerie are shown with a black square and the ones that are under the responsibility of the police are shown with a grey cross. Each panel represents a map of Milan (latitude vs. longitude) and each dot represents a robbery. One can see that in day/time combinations that belong to group 1 the gendarmerie cars cover the northwestern part of the city while the police cars cover the rest. In group 2 day/time combinations the gendarmerie covers the northeastern part and in group 3 the southern one. The few outliers are driven by robberies that have been assigned to i) police or gendarmerie cars that are part of smaller offices (*commissariati*) that are distributed across the city, or ii) the mobile forces (*squadra mobile*), or iii) motor bikers that typically operate in criminal hot spots locations (Mastrobuoni).

For the first two years of data (2008-2009) the police did not share any information derived from the predictive policing system. In January 2010, the police started sending information to the gendarmerie, and by the end of the year they had shared 33 classified reports. For this reason I use just the first two years of data (2008-2009) to evaluate predictive policing.[31]

This quasi-experiment satisfies many of the conditions that an ideal experiment would require, and additionally eases concerns regarding the Hawthorne effects, as patrols are not part of a real experiment. The main drawback is that the assignment of predictive policing follows a predetermined pattern and could potentially be predictable by criminals. For this reason it is going to be important to see whether there is evidence of criminals

[31]Staring in 2010 the productivity of the Carabineri does indeed converge to the one of the *Polizia*.

avoiding the more productive police force, and run proper randomization tests to see whether individual characteristic of the criminals or the *modus operandi*, including the loot, predict whether a robbery is assigned to the police force (and to predictive policing).

The summary statistics in Table 1 show that 73 percent of robberies happened in areas that were patrolled by the police force. Given that at any point in time the police force patrols 67 percent (2/3) of the city, there is no evidence that criminals avoid targets that might be covered by predictive policing. This is even more evident in Table 2. There is no evidence that robbers target areas that are patrolled by the *Carabinieri* more frequently, even when distinguishing first robberies (where no predictive policing can be used) from subsequent ones.

Table 3 performs a balance test, comparing all the observable characteristics of the robbers and of the robberies depending on whether the Polizia or the Carabinieri where covering the area. The single most striking difference is in the likelihood of clearing a robbery. The amount stolen, which is a measure of the ability of robbers (see Mastrobuoni, 2011) does not show any differences. One cannot even reject the hypothesis that for the two forces the whole distribution of the loot, shown in Figure 6, is the same.

There are a few variables for which Table 3 measures small but significant differences between the two forces: year, day of the week, pharmacies and other businesses. It will be shown that controlling for these small differences does not matter. It also does not matter for identification, as it is enough to assume that any differences in the productivity of the police forces or in the characteristics of robberies are constant across first and subsequent robberies.

Finally, if robbers knew about these differences there should be relatively more robberies that fall under the responsibility of the gendarmerie than of the police, especially past the first robbery (when predictive policing might potentially aid the investigation). Table 2 showed this not to be the case. Subsequent robberies, which are the ones for

which the policing software predicts potential targets, time, etc., are not more likely to fall under the responsibility of the gendarmerie.

4.2 Simple Difference and Difference-in-Difference

Table 4 shows the clearance rates by year and by police forces, separating robberies between first and subsequent events of a series (a series is a set of robberies that either have some individuals, or a unique vehicle, or weapon in common). As mentioned before, the reason for separating first and subsequent offenses is that one would not expect predictive policing to work without having previously gathered the data; indeed, overall there are few differences in clearance rates between police and gendarmerie for first events in a series. For subsequent events, instead, clearance rates are much larger for the police than for the gendarmerie. That such differences are significantly different from zero can be seen in the regression Table 5.

I model clearances using a linear probability model, where the dummy variable is equal to one when the n-th robbery within a series i is cleared before the next robbery happens:

$$C_{i,n} = \alpha + \delta_k Police_{i,n} + \gamma' X_{i,n} + \epsilon_{i,n} \,;\, k = 1(n > 1). \tag{1}$$

The coefficient δ_k on *Police Intervention* measures the simple difference in clearance rates between the *Polizia* and the *Carabinieri*. Columns 1 and 2 restricts the analysis to first robberies ($n = 1$), columns 3 and 4 to subsequent ones ($n > 1$). The difference in clearance rates is basically 0 among first robberies and is equal to 10 percentage points (significant at the 1 percent level with standard errors clustered at the series level) for subsequent ones.

Consistent with the quasi-random assignment of police forces, controlling for additional regressors listed at the bottom of the table (columns 2 and 4) leaves the coefficients almost unchanged. Relative to the gendarmerie these results mean that the police officers are

almost 3 times more likely to solve subsequent robberies compared to the gendarmerie officers. If this difference was driven by underlying differences in productivity, e.g. having a more widespread control over the city (2 out of 3 areas), or, possibly, more efficient police officers, one would expect to find a similar difference among first robberies.

There is an additional difference one can exploit. To avoid that the victims' post-traumatic stress might induce a recall bias the Milan police waits 24 hours before collecting the data. This means that when robbers perform two robberies in the same day, patrols are not going to have an updated version of the prediction for the second robbery (keeping in mind that the predictions for the first offense did not work). Given that one is conditioning on a series of robberies where predictive policing did not work for the first robbery, this coefficient measures an upper bound of the effect of predictive policing.[32]

Columns 5 and 6 of Table 5 show that the *Polizia*'s productivity is considerably larger when robbers do not perform their second robbery during the same day of the first robbery. For the gendarmerie no such difference emerges (columns 7 and 8).

In columns 1 and 2 of Table 6 I combine regressions based on Table 5 in a difference-in-difference setup. Comparing *Polizia* versus *Carabineri* for first and subsequent robberies the effect of predictive policing is equal to almost 8 percentage points (Column 1):

$$C_{i,n} = \alpha + \delta_0 Police_{i,n} + \delta_1 1(n > 1) + \delta_2 Police_{i,n} \times 1(n > 1) + \epsilon_{i,n}. \tag{2}$$

In Column 2 I restrict the analysis to subsequent robberies, exploiting differences between same-day and different-day robberies, for *Polizia* and *Carabineri*.[33] The estimated effect is larger, but one needs to keep in mind, as noted earlier, that this strategy estimates a upper bound of the productivity effects of predictive policing.

[32]The estimated equation is similar to Equation 1, where the binary variable 1(different day robbery) is used instead of $1(n > 1)$.

[33]The estimated equation is similar to Equation 2, where the binary variable 1(different day robbery) is used instead of $1(n > 1)$.

Not only should we expect there to be a difference between first and subsequent robberies, but as the police force keeps on collecting information about the robbers, the difference in productivity should also increase. Columns 3 and 4 presents difference-in-difference estimates where the difference is allowed to increase or decrease as a function of the number of robberies performed by the robbers

$$C_{i,n} = \alpha + \delta_0 Police_{i,n} + \delta_1 n + \delta_2 Police_{i,n} \times n + \epsilon_{i,n}. \tag{3}$$

Column 3 shows that when there is a *Police Intervention* the likelihood of clearing a case increases by 0.9 percentage points (more than 10 percent) for each additional robbery (*Number of the series*) the predictive policing software can analyze. The estimate of δ_2 based on same-day/different-day differences is slightly larger, but is subject to the previous caveat.

It is also interesting to notice that for the gendarmerie the coefficients on "Subsequent robberies" and on the "Number of the series" are negative, indicating that either due to selection or learning successful robbers become more and more difficult to arrest. Predictive policing seems to counteract these forces.

Overall there is strong evidence that predictive policing leads to a large increase in the likelihood of solving crimes. All the evidence presented so far points toward a large causal effect of predictive policing on the productivity of police forces. The natural follow-up is to uncover the mechanism through which predictive policing works. For predictive policing to work criminals need to show some persistence in behavior; if such persistence was not visible in the data, predictive policing would hardly be able to explain the large differences in clearance rates.

5 Why Does Predictive Policing Work? Persistence

If the purpose of predictive policing is to optimize police patrolling (delivering a list of potential targets) the two main predictions are about time and location of a robbery.

Several mechanisms can rationalize the predictability of robbers, like, for example, superior information about targets, learning through experience, time constraints (legitimate work, darkness, etc.), or liquidity constraints. Robbers might thus choose to operate in certain parts of the city, against certain businesses, and even in certain times of the day, of the week, and even at regular intervals for completely rational reasons.

Here I test for persistence using all years of data and all the variables that I have been given that could potentially be exploited by a predictive policing software. The easiest way to show persistence in the choice of the location of a robbery is to plot these for each group of robbers. Figure 7 shows the distribution of locations (by latitude and longitude) for groups of robbers with a total of at least 15 robberies. While there is considerable heterogeneity in the amount of clustering, robbers do appear to restrict their activities based on geography. It is also easy to show that the variance in longitude and latitude within groups of robbers is considerably smaller than the variance between groups.

In order to measure persistence I use information that was available to the police before a given robbery. In case of discrete variables (D) a criminal group i shows persistence when some chosen *modus operandi* are identical to the previous most frequent (modal) *modus operandi* ($Persistence(D_{i,t}) = 1(D_{i,t} = mode(D_{i,t-1}, ..., D_{i,0}))$. For example, if most of the first 5 targets were banks, I compute the likelihood that the subsequent target is a bank. Whenever there is more than one mode I randomly select just one. If there was no persistence such likelihood would equal the marginal distribution of business types.

Figure 8 shows that the marginal distributions are several orders of magnitude smaller than the likelihood that a group of robbers targets the type of business they have been

targeting most often in the past.

Figure 9 shows that a very similar pattern emerges when one classifies the time at which a robbery is done into 60-minute periods (the length of the period does not matter). Robbers who are used to target businesses, for example, between 1 and 2 pm (13 in the figure) are very likely to do so again. There is less evidence of persistence in the afternoon.

Finally, Figure 10 shows that there is some persistence in the chosen day of the week, but only Sunday and Monday, and a little bit on Friday and Saturday. Robbers do not seem to develop the habit of robbing businesses on Tuesday, Wednesday, or Thursday.

Days of the month shown little persistence, which is not very surprising given that robbers often operate several times each month. But an additional variable that might signal when the next robbery is going to take place is the time between one robbery and the next. Figure 11 shows that among offenders who recidivate, 58 percent recidivate within a week from the last offense (10 percent recidivate the same day), and that those whose modal recidivism is within a week have a probability that is slightly larger to recidivate again within a week. Otherwise there is little evidence of predictability.

When continuous variables (X) are used to measure persistence I compute their mean absolute deviation from the mean using only past and present data ($Persistence(X_{i,t}) = -t^{-1}\sum_{\tau\leq t}|X_{i,\tau} - \overline{X}_{i,t}|$, where $\overline{X}_{i,t} = t^{-1}\sum_{\tau\leq t}X_{i,\tau}$). A larger deviation, for example, in longitude and latitude means that offenders are more mobile and thus exhibit less persistence.

Persistence across one dimension might clearly be correlated with persistence across other dimensions. Table 7 shows evidence that robbers who often select their modal hour of the day tend to select their modal type of business.

The single most important factors that appear to be predictable based on the the graphical analyses are the time of the day, the type of victimized business, the distance between robberies, and the time between a robbery and the next. In Table 8 I regress each

of these factors on a measure of experience (the *Number of the series*), of success (whether in the previous robbery the *Previous loot was larger than average for business*), as well as on several characteristics of the robbers. The purpose is to see what is associated with persistence. Given the lack of exogenous variation these regressions are of descriptive nature. Robbers who have performed a successful robbery (a robbery whose loot was larger than average for that type of business) is 3 percentage points more likely to chose the same hour for his/her next robbery. Given that on average 1 in 10 chooses the same hours this represents a 30 percent increase. Persistence in the time of day increases also with experience. Every additional robbery is associated with an additional 0.003 more persistence in time, or 3 percent.

Robbers whose previous loot was higher than average seem to be more likely to select the same type of business. The log-distance between victims as well as persistence in weeks between robberies are associated with experience but not with the success level of the previous robbery. As robbers get more experienced they wait less (though more likely within the same week) but move more, possibly to find new targets.

Most individual characteristics do not predict persistence with the notable exception of the number of robbers involved. Each additional robber makes the robbery more unpredictable: the persistence in time props by 27 percent, the one about distance by 28 percent, and the one about time between the robberies by 20 percent.

6 Policy Implications

Clearing a robbery is synonymous with arresting at least one robber.[34] Based on data collected by the police the 31 series that were cleared in 2008 led to a total of 203 years in jail and the 39 series cleared in 2009 to 217 years in jail. Given that the average number of

[34]According to the police a few times they waited to make the arrest of identified perpetrators only to gather additional evidence.

robbers per robbery is 1.5, about 100 robbers were arrested, and their average conviction is close to 4 years of jail time. Of these robbers, only one was found not guilty and 4 were given alternative sanctions to prison time.

After their first robbery about 30 percent of robbers organize a second robbery, and after that almost all keep on robbing business until they get arrested (see Mastrobuoni, for evidence on this "life" table of robberies), differences in clearance rates lead to differences in the expected number of robberies these criminal groups are able to organize before ending up in jail. Since the police and the gendarmerie share these incapacitation effects (there is quasi-random assignment of crimes to the two police forces), such effects cannot be measured directly. But one can use the differences in clearance rates and some simple algebra to retrieve such effects.

Using the more conservative treatment effect (the 8 percentage points difference in difference in clearance rates between the police and the gendarmerie for subsequent and first offenses), the expected number or robberies each group commits, which is equal to $\sum_{\tau=1}^{\infty}(r_t(1-c_t))^{\tau}$ (r is the likelihood of reiterating a robbery, and c is the clearance rate) drops from about 18 to 6.6.

Since there are about 255 successful first time robbers a year and about one-third re-offend, the reduction of 11 robberies per series leads to a total reduction of 935 robberies (in the long run deterrence might lead to even larger reductions). Multiplying such number by the average haul (€2,800) the direct costs that are prevented by the use of predictive policing are close to €2.5 million, or more than about US$ 3 million.[35]

In order to evaluate the cost and benefits of predictive policing one would also need to take into account the increased cost from incarcerating arrested criminals and the cost of investing in such IT. Since most of these robbers would end up in jail anyway, or in

[35]Indirect costs are likely to be an order of magnitude larger that the direct costs Cook (2009). Moreover, public concern about crime is to a large extent a concern about robbery, and might lead to a "secondary mischief," (Bentham, 1879), constraining choices about where to live, work, shop, and go out to dinner (**?** find evidence of such indirect costs.)

other words, since $(1-c_t)^{\tau}$ converges to 0 reasonably quickly for clearance rates that are close to 10 percent, the counterfactual cost is likely to be similar.

The most "costly" incarcerations would be related to those robbers who perform a robbery but would have stopped committing crimes anyway. Arresting these robbers just before they quit would generate an inefficient but possibly just incapacitation (retribution and deterrence would still speak in favor of the arrest). There are only very few such robbers. Of the 100 recidivists that move from the first to the second robbery the predictive policing would lead to an additional arrest of 10 individuals at or after their second robbery and before the third. 18 percent of these, or about 2, would have quit robbing commercial businesses in Milan (though they might have moved to other criminal enterprises or to other cities). Past the second robbery basically offenders reiterate the crime until incapacitation.

At an average cost of €50,000 per inmate, in four years the 2 "inefficient" incarcerations would generate an additional €400,000 in public spending (Barbarino and Mastrobuoni, 2014), clearly an order of magnitude lower than the benefits.[36] The labor cost of the three fulltime police officers who collect the data and predict the crimes is also below €100,000 a year. The investments in capital (an office, computers, monitors, etc) would hardly be above a few thousand euros a year.

Additional cost and benefits are related to how the additional information collected through predictive policing helps the prosecutors to build a case in court. Unfortunately there are no data (e.g. post-incarceration recidivism of convicted robbers) to evaluate such cost and benefits, though they are arguably smaller in magnitude than the direct cost of crimes, and would hardly turn around the cost/benefit findings.

[36]Despite being "inefficient" such incarcerations would be targeting guilty robbers. And yet, victimizations as well as incarcerations generate pain and suffering which I do not attempt to quantify, as both are extremely hard to measure.

7 Conclusions and Policy Implications

This study used the quasi-random allocation of two almost identical police forces to crimes, to test whether differences in police productivity can be attributed to the availability of advanced Information Technology. The differences are striking, but only among subsequent robberies, thus only once data to be analyzed become available. The second part of the study shows that robbers are indeed predictable based on their past actions.

A rough cost/benefit analysis suggests that micro-predictive policing represents a highly cost efficient IT investment. Related to the cost/benefit analysis it is worth highlighting that, because of its inherent nature, the micro-predictive policing IT innovation helps securing the most prolific criminals. The more prolific they are, the more data can be collected and the more productive the *Polizia* becomes (compared to the *Carabinieri*). Since these criminals tend to be the most socially harmful, predictive policing leads to more selective incarcerations.

Overall, the first quasi-experimental evaluation of micro-based predictive policing against commercial robberies suggests that such IT investments can be highly effective.

References

Daron Acemoglu, Philippe Aghion, Claire Lelarge, John Van Reenen, and Fabrizio Zilibotti. Technology, information, and the decentralization of the firm. *The Quarterly Journal of Economics*, 122(4):1759–1799, November 2007.

Agatha Christie. *The A.B.C. Murders*. Collins Crime Club, 1936.

Joshua Angrist and Victor Lavy. New evidence on classroom computers and pupil learning*. *The Economic Journal*, 112(482):735–765, 2002.

Susan Athey and Scott Stern. The impact of information technology on emergency health care outcomes. *Rand Journal of Economics*, 33(3):399–432, 2002.

David H. Autor, Lawrence F. Katz, and Alan B. Krueger. Computing inequality: Have computers changed the labor market? *The Quarterly Journal of Economics*, 113(4): 1169–1213, November 1998.

Marzio Barbagli and Asher Colombo. *Rapporto sulla criminalità e la sicurezza in Italia*. Gruppo24Ore, 2011.

Alessandro Barbarino and Giovanni Mastrobuoni. The Incapacitation Effect of Incarceration: Evidence from Several Italian Collective Pardons. *American Economic Journal: Economic Policy*, 6(1):1–37, 2014.

Ann P. Bartel, Casey Ichniowski, and Kathryn L. Shaw. How does information technology really affect productivity? plant-level comparisons of product innovation, process improvement and worker skills. *The Quarterly Journal of Economics*, 4(122), 2007.

Jeremy Bentham. *An introduction to the principles of morals and legislation*. The Clarendon Press, 1879.

Eli Berman, John Bound, and Zvi Griliches. Changes in the demand for skilled labor within u.s. manufacturing: Evidence from the annual survey of manufactures. *The Quarterly Journal of Economics*, 109(2):367–97, May 1994.

Sandra E. Black and Lisa M. Lynch. How to compete: The impact of workplace practices and information technology on productivity. *The Review of Economics and Statistics*, 83(3):434–445, August 2001.

Anthony A. Braga. The effects of hot spots policing on crime. *Annals of the American Academy of Political and Social Science*, 578:pp. 104–125, 2001.

Timothy F. Bresnahan, Erik Brynjolfsson, and Lorin M. Hitt. Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *The Quarterly Journal of Economics*, 117(1):339–376, February 2002.

Erik Brynjolfsson and Lorin M. Hitt. Beyond computation: Information technology, organizational transformation and business performance. *Journal of Economic Perspectives*, 14(4):23–48, Fall 2000.

Paolo Buonanno and Giovanni Mastrobuoni. Police and crime: Evidence from dictated delays in centralized police hiring. Technical report, 2011. mimeo.

Ronald Clarke. *Situational crime prevention*. Criminal Justice Press, 1997.

Ronald V Clarke and David Weisburd. Diffusion of crime control benefits: Observations on the reverse of displacement. *Crime prevention studies*, 2:165–184, 1994.

Jacqueline Cohen and Jens Ludwig. Policing crime guns. In *Evaluating Gun Policy: Effects on Crime and Violence*, page 217. Brookings Institution Press, 2003.

Philip Cook, Jens Ludwig, and Justin McCrary. Economical crime control. In *Controlling*

Crime: Strategies and Tradeoffs, NBER Books, pages 331–363. National Bureau of Economic Research, Inc, 2011.

Phillip Cook. Robbery. In Michael Tonry, editor, *Oxford Handbook on Crime and Public Policy*. Oxford University Press, 2009.

Hope Corman and H. Naci Mocan. A Time-Series Analysis of Crime, Deterrence, and Drug Abuse in New York City. *American Economic Review*, 90(3):584–604, June 2000.

Paul A. David. The dynamo and the computer: An historical perspective on the modern productivity paradox. *American Economic Review*, 80(2):355–61, May 1990.

Rafael Di Tella and Ernesto Schargrodsky. Do police reduce crime? Estimates using the allocation of police forces after a terrorist attack. *The American Economic Review*, 94 (1):115–133, 2004.

Mark Doms, Timothy Dunne, and Kenneth R Troske. Workers, wages, and technology. *The Quarterly Journal of Economics*, 112(1):253–90, February 1997.

Mirko Draca, Stephen Machin, and Robert Witt. Panic on the Streets of London: Police, Crime and the July 2005 Terror Attacks. *American Economic Review*, 101(5):2157–81, 2011.

Steven N. Durlauf, Salvador Navarro, and David A. Rivers. Understanding aggregate crime regressions. *Journal of Econometrics*, 158(2):306 – 317, 2010. ISSN 0304-4076. Specification Analysis in Honor of Phoebus J. Dhrymes.

The Economist. The Aftershocks of Crime. *The Economist*, October 21 2010.

William N. Evans and Emily G. Owens. COPS and Crime. *Journal of Public Economics*, 91(1-2):181–201, 2007. ISSN 0047-2727.

Luis Garicano and Paul Heaton. Information technology, organization, and productivity in the public sector: Evidence from police departments. *Journal of Labor Economics*, 28(1):167–201, 01 2010.

Austan Goolsbee and Jonathan Guryan. The impact of internet subsidies in public schools. *The Review of Economics and Statistics*, 88(2):336–347, May 2006.

Lev Grossman, Cleo Brock-Abraham, Nick Carbone, Eric Dodds, Jeffrey Kluger, Alice Park, Nate Rawlings, Claire Suddath, Feifei Sun, Mark Thomson, Bryan Walsh, and Kayla Webley. The 50 best inventions. *Time Magazine*, November 28 2011.

Jonathan Klick and Alexander Tabarrok. Using terror alert levels to estimate the effect of police on crime. *Journal of Law & Economics*, 48(1):267–79, April 2005.

Steven D Levitt. Using electoral cycles in police hiring to estimate the effect of police on crime. *American Economic Review*, 87(3):270–90, June 1997.

Steven D. Levitt. Understanding why crime fell in the 1990s: Four factors that explain the decline and six that do not. *Journal of Economic Perspectives*, 18(1):163–190, September 2004.

Steven D. Levitt and T. Miles. Empirical Study of Criminal Punishment. *The Handbook of Law and Economics*, 2004.

Stephen Machin and Olivier Marie. Crime and police resources: The street crime initiative. *Journal of the European Economic Association*, 9(4):678–701, 2011. ISSN 1542-4774.

Alexandre Mas. Pay, reference points, and police performance. *The Quarterly Journal of Economics*, 71(3):783–821, 2006.

Giovanni Mastrobuoni. Police and Clearance Rates: Evidence from Recurrent Redeployment Within a City. mimeo.

Giovanni Mastrobuoni. Optimal criminal behavior and the disutility of jail: Theory and evidence on bank robberies. Carlo Alberto Notebooks 220, Collegio Carlo Alberto, 2011.

Giovanni Mastrobuoni and Emily Owens. Criminal careers and criminal firms. 2014.

Justin McCrary. Dynamic perspectives on crime1. *Handbook on the Economics of Crime*, page 82, 2010.

George O. Mohler, Martin B. Short, Jeffrey Brantingham, Frederick P. Schoenberg, and George E. Tita. Self-exciting Point Process Modeling of Crime. *Journal of the American Statistical Association*, 106(493):100–108, 2011.

Emily G. Owens. Cops and cuffs. In Philip Cook, Stephen Machin, Olivier Marie, and Giovanni Mastrobouni, editors, *Lessons from the Economics of Crime: What Works in Reducing Offending?* Cambridge: MIT Press, 2014.

Beth Pearsall. Predictive policing: The future of law enforcement? *National Institute of Justice*, 2010.

Lawrence W. Sherman and David Weisburd. General deterrent effects of police patrol in crime "hot spots": A randomized, controlled trial. *Justice Quarterly*, 12(4):625–648, 1995.

Lawrence W. Sherman, Patrick R. Gartin, and Michael E. Buerger. Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology*, 27(1):27–56, 1989.

Kevin J. Stiroh. Information technology and the u.s. productivity revival: What do the industry data say? *American Economic Review*, 92(5):1559–1576, December 2002.

David Weisburd and John E Eck. What can police do to reduce crime, disorder, and fear? *The Annals of the American Academy of Political and Social Science*, 593(1): 42–65, 2004.

David Weisburd and Lorraine Green. Policing drug hot spots: The jersey city drug market analysis experiment. *Justice Quarterly*, 12(4):711–735, 1995.

David Weisburd, Stephen D. Mastrofski, Ann McNally, Greenspan Rosann, and James J. Willis. Reforming to Preserve: Compstat and Strategic Problem Solving in American Policing. *Criminology & Public Policy*, 2(3):421–456, 2003.

David Weisburd, Laura A. Wyckoff, Justin Ready, John E. Eck, Joshua C. Hinkle, and Frank Gajewski. Does crime just move around the corner? a controlled study of spatial displacement and diffusion of crime control benefits. *Criminology*, 44(3):549–592, 2006.

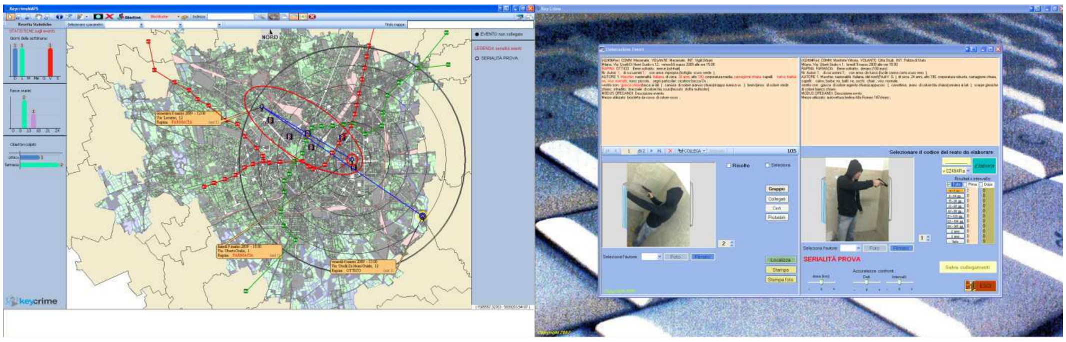

Figure 1: Comparison of Events

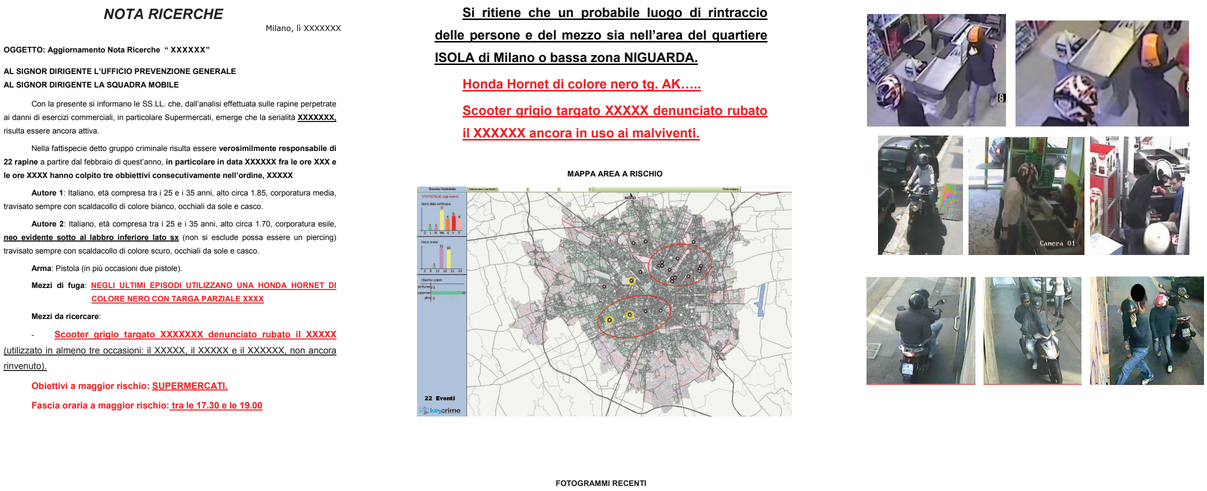

Figure 2: Instructions for Police Patrols

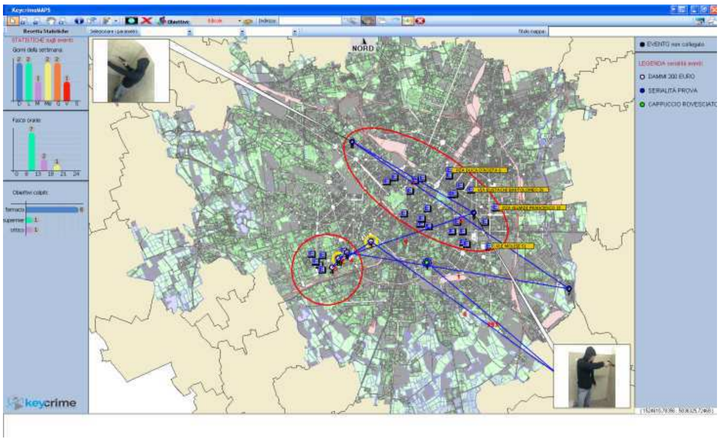

Figure 3: Predicted Targets

Notes: Small blue dots indicate past victims, red circles indicate potential targeted areas, while the little blue squares indicate potential victims.

Figure 4: Gendarmerie and Police

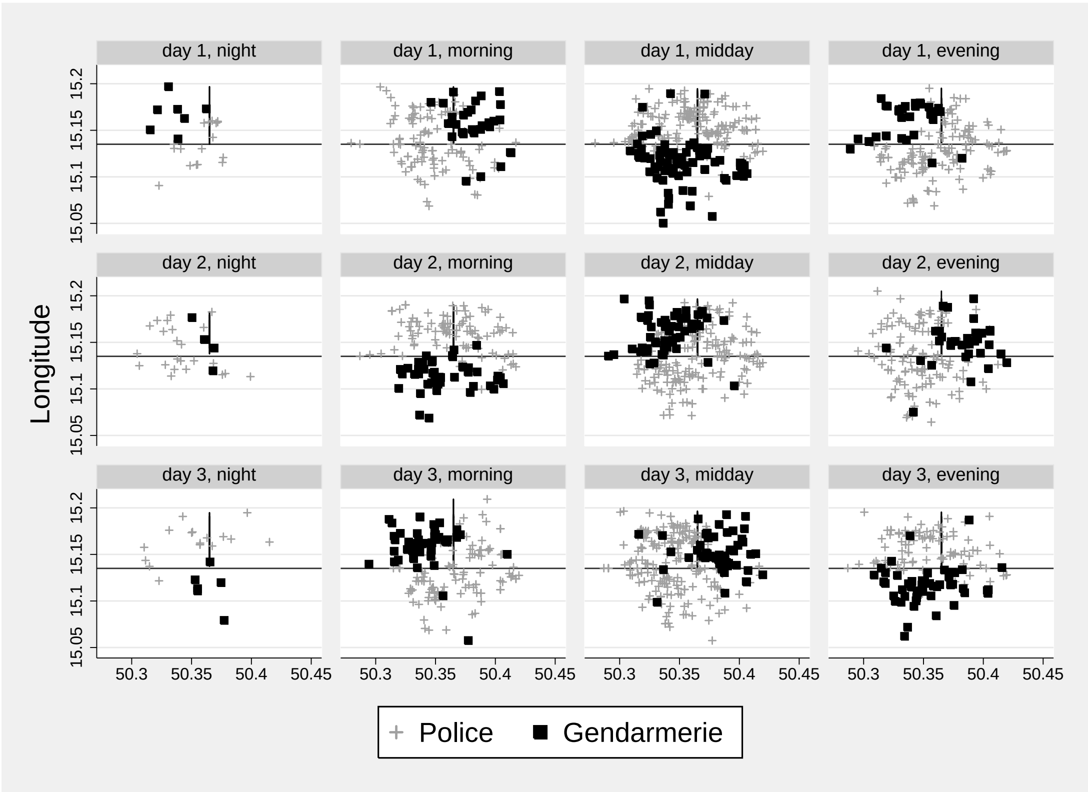

Figure 5: Geographic Distribution of Robberies by Group

Notes: Groups are defined based on the exact day and time of a robbery.

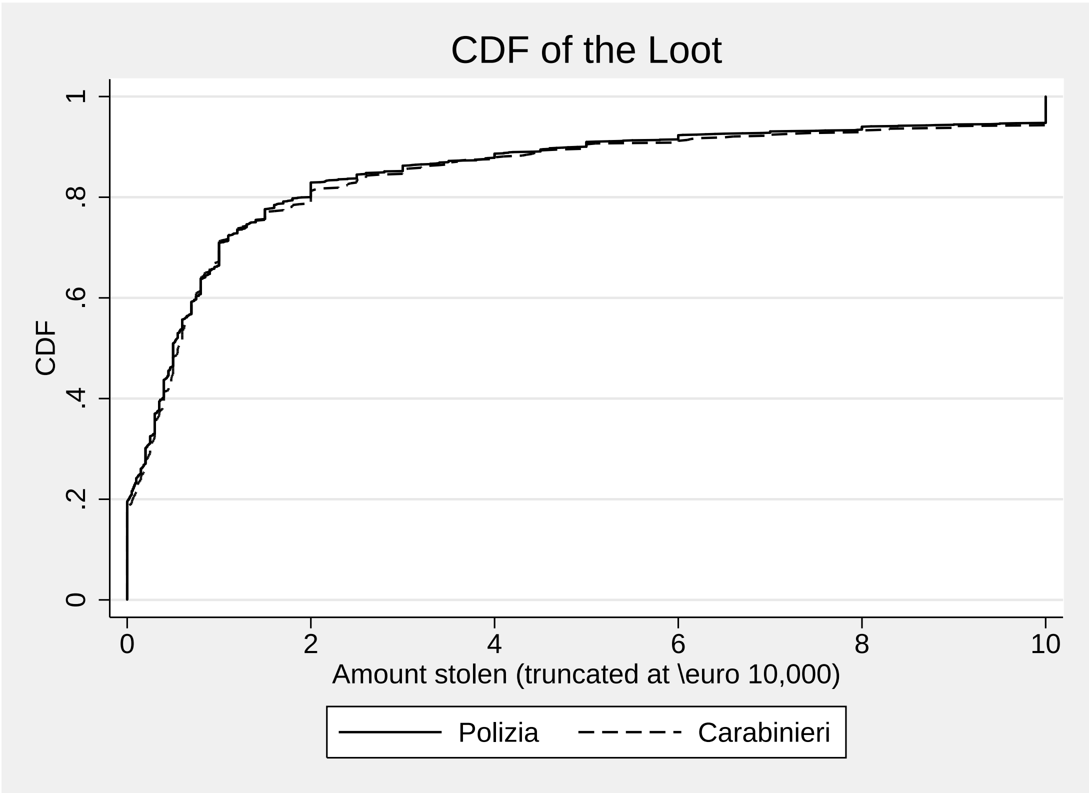

Figure 6: Cumulative Distribution of the Loot

Notes: The loot is expressed in €1,000 and is truncated at 10,000. The Kolmogorov-Smirnov test for equality of distribution functions cannot reject that null that the non-truncated distributions are the same.

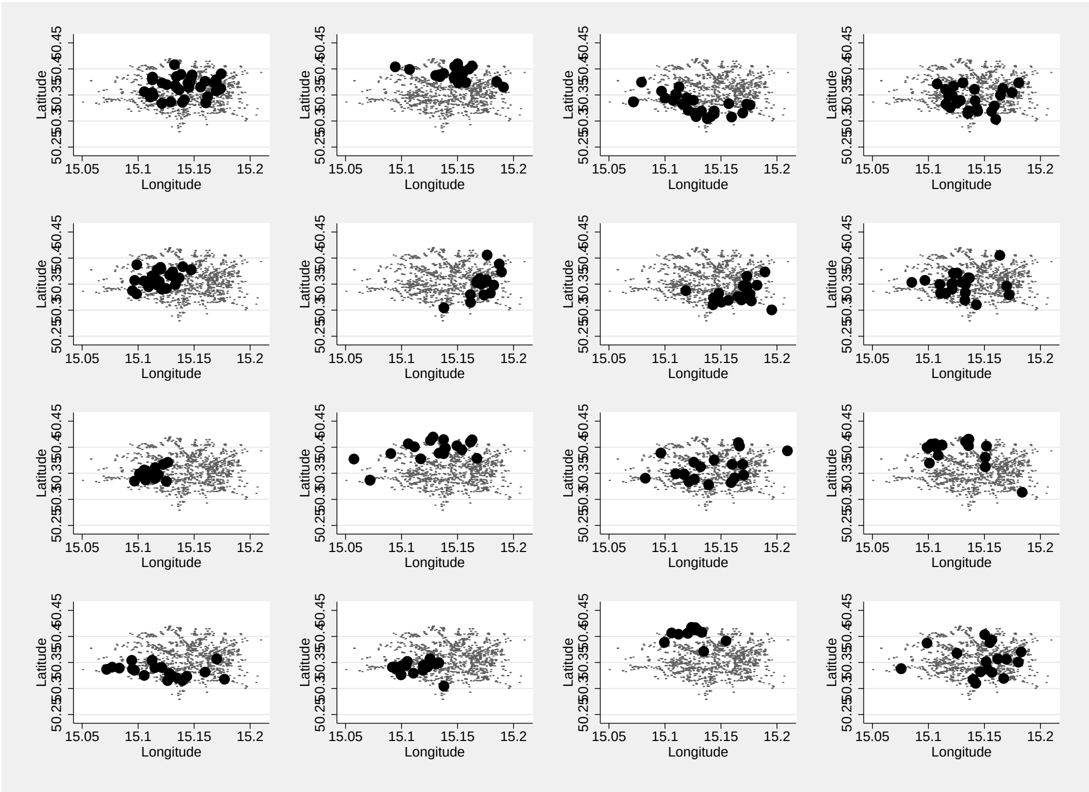

Figure 7: Geographic Distribution of Robberies by Criminal Group

Notes: The plots are restricted to those groups who performed at least 15 robberies. Surveillance camera are used to identify the same offenders across robberies. For each group robberies are labeled sequentially.

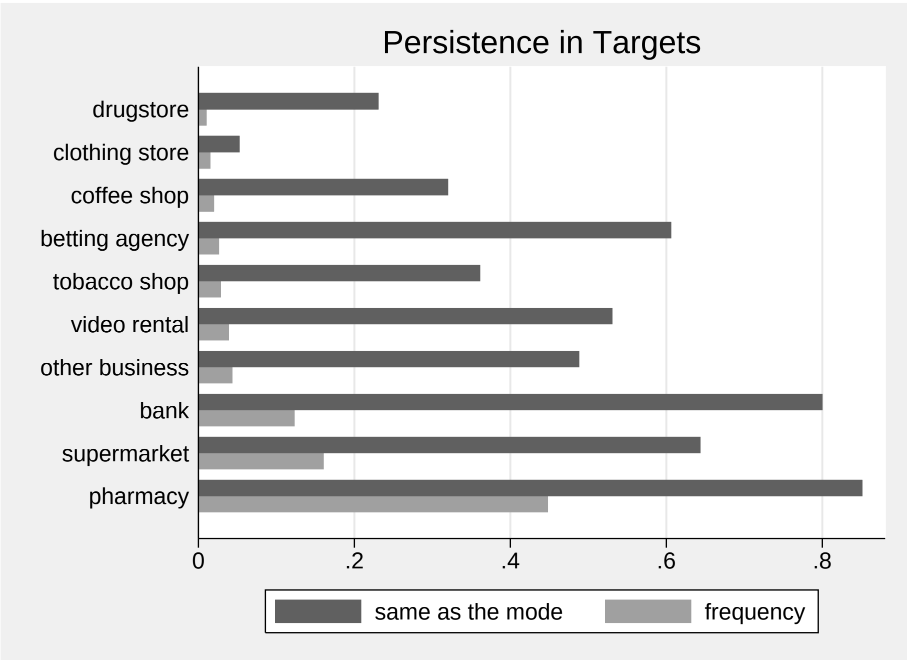

Figure 8: Persistence in Targets

Notes: The dark bar shows the fraction of robbers who select a type of business that is equal to the modal type of business they have been selecting before that robbery. The grey bar represent the simple frequencies. There are 27 different types of business and the figure shows only those businesses that represent at least 1 percent of targets.

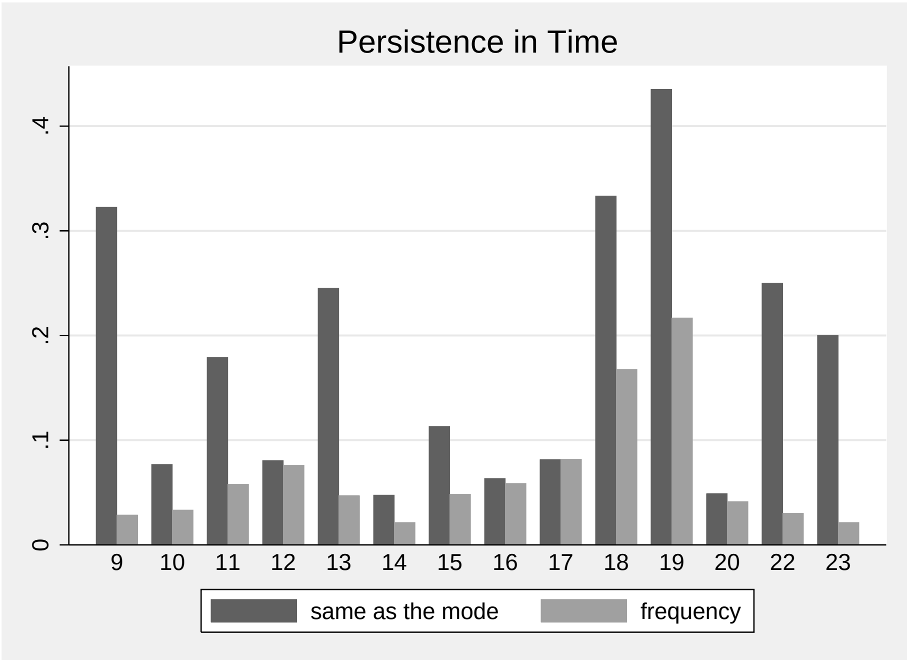

Figure 9: Persistence in Time

Notes: The dark bar shows the fraction of robbers who select an hour that is equal to the modal hour they have been selecting before that robbery. The grey bar represent the simple frequencies. The figure shows only those hours that represent at least 2 percent of the data.

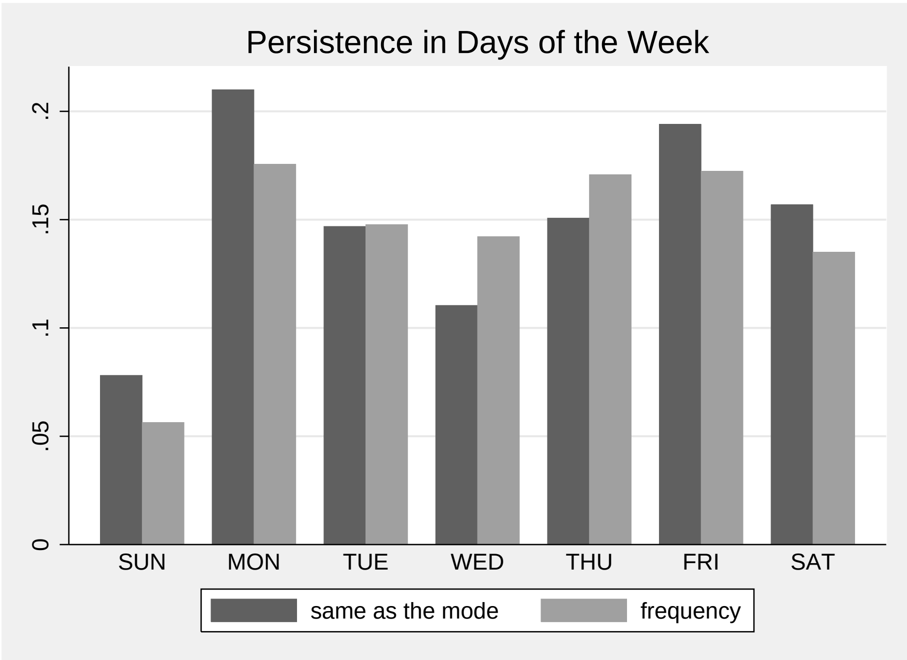

Figure 10: Persistence in Days of the Week

Notes: The dark bar shows the fraction of robbers who select a day of the week that is equal to the modal day of the week they have been selecting before that robbery. The grey bar represent the simple frequencies.

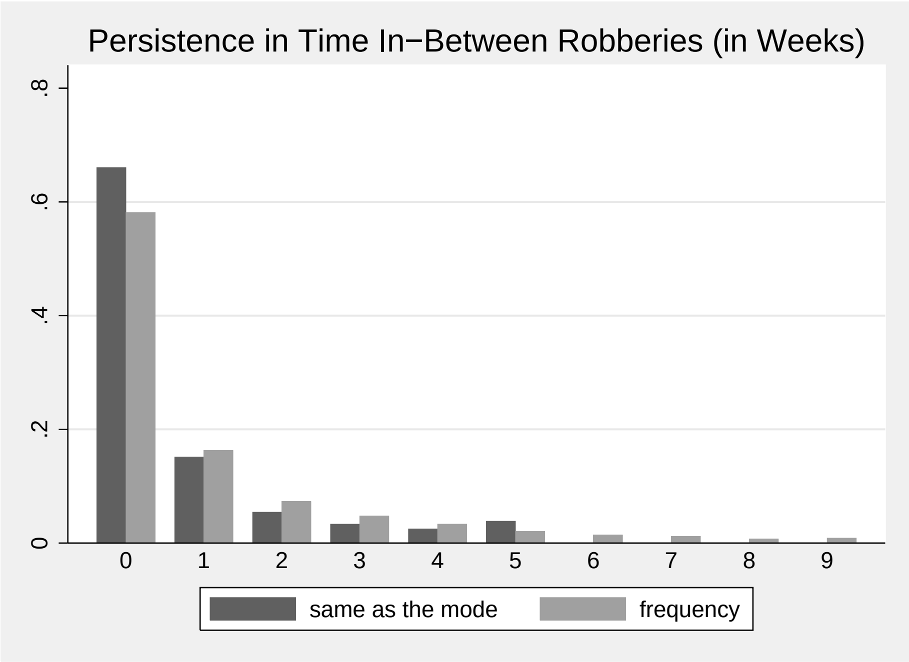

Figure 11: Persistence in Time In-Between Robberies

Notes: The dark bar shows the fraction of robbers who select a time in-between robberies (in weeks, truncated at 9 weeks) that is equal to the modal time in-between robberies they have been selecting before that robbery. The grey bar represent the simple frequencies.

Table 1: Summary statistics

Variable	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
	Full Sample (2008-2011)				Restricted Sample (2008-2009)			
Cleared robbery	0.14	0.35	0	1	0.13	0.34	0	1
Cleared series	0.45	0.50	0	1	0.44	0.50	0	1
Number of the series	5.10	6.88	1	84	4.20	5.53	1	84
Police 0/1	0.73	0.44	0	1	0.74	0.44	0	1
Days between subsequent	16.80	46.43	0	555	14.48	43.47	0	555
Subsequent robberies	0.58	0.49	0	1	0.54	0.50	0	1
North-Western area	0.35	0.48	0	1	0.38	0.48	0	1
North-Eastern area	0.22	0.41	0	1	0.19	0.39	0	1
Year	2009.24	1.02	2008	2011	2008.47	0.50	2008	2009
Month	5.88	3.71	1	12	6.20	3.75	1	12
Day of the month	15.60	8.86	1	31	15.74	8.97	1	31
Day of the week	3.24	1.83	0	6	3.19	1.82	0	6
Daylight	0.59	0.49	0	1	0.57	0.49	0	1
Average age	26.57	12.47	0	68	26.14	13.10	0	68
Amount stolen in euros	2.86	11.18	0	206	2.11	7.90	0	100
Firearm 0/1	0.23	0.42	0	1	0.21	0.41	0	1
At least one knife, but	0.09	0.29	0	1	0.09	0.28	0	1
Some Italian	0.79	0.41	0	1	0.77	0.42	0	1
Different nationalities	0.14	0.35	0	1	0.12	0.32	0	1
Number of robbers	1.57	0.72	1	7	1.51	0.68	1	5
Obs		2167				1255		

Table 2: Distribution of Robberies across Forces

	First Robbery		Subsequent Robberies	
	Carabinieri	Polizia	Carabinieri	Polizia
2008	25.8	74.2	22.62	77.38
2009	30.51	69.49	29.05	70.95
2010	29.46	70.54	30.42	69.58
2011	22.22	77.78	23.25	76.75
Total	27.54	72.46	26.77	73.23

Table 3: Balance Test

	Police		Gendarmerie		Police-Gendarmerie	
	Average	SE	Average	SE	Average	SE
Cleared robbery	0.149	0.012	0.093	0.017	0.056	0.020***
Cleared series	0.450	0.039	0.429	0.049	0.022	0.036
Number of the series	4.203	0.464	4.202	0.725	0.001	0.524
Days between subsequent robberies	14.978	2.271	12.887	2.542	2.091	3.013
Subsequent robberies	0.550	0.031	0.500	0.041	0.050	0.036
Shift change 0/1	0.160	0.014	0.146	0.020	0.014	0.023
Shift	3.055	0.041	2.963	0.056	0.092	0.057
North-Western area	0.375	0.027	0.376	0.037	-0.001	0.033
North-Eastern area	0.188	0.019	0.199	0.027	-0.011	0.026
Year	2008.450	0.036	2008.534	0.042	-0.084	0.035**
Month	6.151	0.246	6.351	0.315	-0.200	0.261
Day of the month	15.868	0.363	15.357	0.492	0.511	0.585
Sunday	0.054	0.008	0.071	0.014	-0.018	0.015
Monday	0.163	0.013	0.233	0.023	-0.070	0.025***
Tuesday	0.159	0.012	0.137	0.019	0.022	0.023
Wednesday	0.143	0.011	0.155	0.019	-0.013	0.022
Thursday	0.189	0.013	0.127	0.018	0.061	0.022***
Friday	0.167	0.013	0.149	0.024	0.018	0.028
Saturday	0.126	0.012	0.127	0.020	-0.001	0.023
Daylight	0.564	0.025	0.602	0.033	-0.039	0.034
Average age	26.080	0.655	26.308	1.205	-0.229	1.132
Amount stolen in euros	1.921	0.264	2.664	0.591	-0.743	0.536
Firearm 0/1	0.198	0.023	0.233	0.033	-0.035	0.029
At least one knife, but no firearm	0.084	0.016	0.090	0.020	-0.006	0.018
Some Italian	0.778	0.023	0.752	0.033	0.027	0.030
Different nationalities	0.120	0.011	0.112	0.019	0.008	0.020
Number of robbers	1.514	0.038	1.516	0.052	-0.001	0.046
Pharmacy	0.356	0.036	0.422	0.043	-0.067	0.034**
Other business	0.160	0.017	0.118	0.020	0.042	0.023*
Supermarket	0.152	0.021	0.165	0.025	-0.012	0.026
Bank	0.073	0.019	0.096	0.022	-0.023	0.021
Video rental	0.033	0.014	0.053	0.018	-0.020	0.013
Tobacco shop	0.025	0.007	0.016	0.010	0.009	0.010

Notes: Years 2008/2009. Standard errors are clustered by series. For the last two columns only: *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Clearance Rates by Year and Police Force

	First event		Subsequent events	
	Gendarmerie	Police	Gendarmerie	Police
2008	0.124	0.160	0.049	0.121
	(0.331)	(0.367)	(0.218)	(0.326)
	89	*256*	*61*	*257*
2009	0.139	0.128	0.060	0.180
	(0.348)	(0.335)	(0.239)	(0.385)
	72	*164*	*100*	*256*
Total	0.130	0.148	0.056	0.150
	(0.338)	(0.355)	(0.230)	(0.358)
	161	*420*	*161*	*513*

Notes: Standard deviations (not errors) are shown in parentheses, the number of observations are shown in italics.

Table 5: Simple Difference Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Clearance rates							
	First robbery		Subsequent robberies		Police in subsequent		Gendarmerie in subsequent	
Police Intervention	0.016	0.018	0.100***	0.087***				
	(0.032)	(0.030)	(0.025)	(0.027)				
Different day					0.122***	0.129***	0.001	0.047
					(0.031)	(0.038)	(0.061)	(0.075)
Constant	0.120***	0.376***	0.074***	-0.066	0.070**	-0.150	0.059	-0.040
	(0.032)	(0.096)	(0.022)	(0.105)	(0.031)	(0.123)	(0.066)	(0.089)
Other Xs		√		√		√		√
Observations	581	581	674	674	510	510	160	160
R-squared	0.001	0.209	0.020	0.101	0.020	0.120	0.000	0.290

Notes: Linear probability models with clustered (by series) standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. All regressions control for a year 2009 fixed effect. The regressions that control for additional regressors contain the following fixed effects: month, day of the week, shift-turnover, morning, evening, and night shift, daylight, Western, North-eastern part of the city, firearm, knife, "some Italian," "different nationalities," pharmacy, other business, supermarket, bank, video rental, tobacco shop. These regressions control also for average age, loot, number of offenders, day of the month, number of the series.

Table 6: Difference in Differences Estimates

	(1)	(2)	(3)	(4)
	The individual robbery has been cleared			
Robberies:	*All*	*Subsequent*	*All*	*Subsequent*
Police Intervention	0.018	-0.022	0.021	0.065*
	(0.032)	(0.057)	(0.026)	(0.035)
Subsequent robberies	-0.078**			
	(0.032)			
Different day robbery		-0.011		0.026
		(0.056)		(0.051)
Number of the series			-0.005***	-0.001
			(0.002)	(0.003)
Number of the series *times* Different day robbery				-0.000
				(0.003)
Police Intervention interacted with:				
Subsequent robberies	**0.078***			
	(0.040)			
Different day robbery		**0.133****		
		(0.064)		
Number of the series			**0.009*****	-0.006**
			(0.003)	(0.003)
N. of the series × Different day robbery				**0.012*****
				(0.004)
Observations	1,255	670	1,255	670
R-squared	0.009	0.030	0.010	0.034

Notes: Linear probability models with clustered (by series) standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. All regressions control for year effects.

Table 7: Correlation in Persistence

	A	B	C	D	E	F	G
A: Selects previous modal hour	1.00						
B: Select previous modal shift	**0.38**	1.00					
C: Current absolute deviation in time	**-0.24**	**-0.33**	1.00				
D: Select previous modal type of business	**0.08**	**0.11**	**-0.16**	1.00			
E: Select previous modal day of the week	0.01	**0.10**	-0.04	0.02	1.00		
F: Select previous modal day of the month	0.02	0.02	0.00	0.01	0.02	1.00	
G: Select previous time between robberies	-0.01	0.05	**0.09**	0.02	-0.04	-0.03	1.00
H: Current absolute deviation in location	-0.01	-0.03	0.04	-0.03	0.00	-0.04	0.02

Notes: Correlation coefficients. The ones in bold are significant at the 10 percent level.

Table 8: Persistence and Success of the Robbery

	(1) Same Hour (0/1)	(2) Same Type of Business (0/1)	(3) Log-distance Between Victims	(4) Same Weeks Between Robberies	(5) Days Between Robberies
Number of the series	**0.003****	0.004	**0.009****	**0.017*****	**-0.776*****
	(0.002)	(0.004)	**(0.004)**	**(0.005)**	**(0.198)**
Loot is larger than average for business	**0.031***	**0.111*****	0.068	0.024	0.268
	(0.019)	**(0.027)**	(0.062)	(0.032)	(2.825)
Average age	0.000	0.003*	0.001	-0.002	0.219*
	(0.001)	(0.002)	(0.003)	(0.002)	(0.118)
Firearm	0.022	-0.110	0.117	-0.133***	1.369
	(0.022)	(0.070)	(0.078)	(0.049)	(4.222)
At least one knife, but no firearm	0.030	0.038	0.123	0.074	-7.867**
	(0.025)	(0.056)	(0.086)	(0.058)	(3.568)
Some Italian	-0.005	-0.010	0.093	0.062	-0.626
	(0.021)	(0.059)	(0.105)	(0.047)	(3.333)
Different nationalities	-0.009	-0.004	-0.054	0.083*	-3.591
	(0.026)	(0.057)	(0.088)	(0.043)	(5.493)
Number of robbers	-0.027*	-0.030	0.213***	-0.081**	7.321*
	(0.015)	(0.050)	(0.052)	(0.034)	(4.264)
Constant	0.101**	0.609***	0.177	0.388***	6.929
	(0.043)	(0.087)	(0.162)	(0.087)	(6.495)
Observations	1,255	1,255	1,201	1,255	1,255
R-squared	0.013	0.038	0.057	0.100	0.026
Mean dep. var.	0.101	0.682	0.734	0.403	16.80

Notes: For the variable “Previous loot was larger than average” the average is based on the business types used in Fig. 8. Standard errors are clustered by series. For the last two columns only: *** p<0.01, ** p<0.05, * p<0.1.