Crime is Terribly Revealing: Information Technology and Police Productivity*

Giovanni Mastrobuoni†

March 2016‡

Abstract

In an unprecedented information technology (IT) revolution in the public service sector, an increasing number of police departments use advanced statistical methods to optimize their patrolling strategies. The most advanced ones are known as "predictive policing" and are capable of predicting individual crimes. An open question is whether predictive policing is also capable of improving the productivity of policing.

I address this question using quasi-random assignment of individual crimes to predictive policing. The adoption of predictive policing leads to a sizable increase in the likelihood that individual crimes are cleared. Detailed information on *individual* crime incidents coupled with offender-level identifiers shed light on the mechanisms behind the productivity improvements. Criminals are shown to follow habits with respect to the type of victims, their location, the time of the offence, and the frequency of offending, especially when their previous actions have brought unexpected criminal proceeds. These habits make their future actions predictable. The benefit-cost ratio of this IT innovation appears to be larger than 10.

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 Phil Cook, Arie Kapteyn, Mark Kleinman, Daniel Nagin, Santiago Oliveros, Emily Owens, David Rivers, and John Von Reenen, as well as seminar participants at the Collegio Carlo Alberto, the Aalto University, the Home Office (UK), the Al Capone workshop, the International Crime and Intelligence Analysis Conference (Manchester, UK), and the NBER Economics of Crime Summer Institute for invaluable comments.

†Department of Economics, University of Essex, Collegio Carlo Alberto and IZA, 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 have shown a dramatic increase in the use of information technology (IT). The purpose is often to predict individual behavior, be it patients, consumers, firms, taxpayers, or criminals. Prominent applications developed by electronic commerce companies, like Google, Amazon, Netflix, etc., predict consumer preferences based on individual browsing and purchase history to generate custom-tailored Internet content;[1] social networking services, like Facebook, use individual friendship networks and posting behavior to predict new connections.

Predictive algorithms are also being developed in the public sector. Data mining techniques to forecast fraudulent behavior, including tax evasion, are spreading across internal revenue services (see Bolton and Hand, 2002, for a review). Focussed policing strategies that are based on the geographical distribution of previous crime incidents are more and more common across law enforcement agencies. For instance, between 1987 and 2003 the proportion of law enforcement agencies that use computers for criminal investigations, dispatch and fleet management, went up respectively from 11, 9 and 7 percent to 59, 58, and 34 percent. Nowadays more than 90 percent of agencies use

[1] These are known as recommendation algorithms (see Adomavicius and Tuzhilin, 2005).

computers to maintain and analyze criminal incidents.[2]

The adoption of IT in the public service sector, which most times is not driven by market forces, may generate huge gains for the public. Yet, without market forces determining what works and what does not is in the hands of analysts and researchers to spot best practice. While a large body of research investigated the relationships between IT, work organization, and productivity,[3] only a few studies show direct evidence about the role of IT in increasing service sector productivity.[4] Moreover, IT investments are often intangible and disproportionately difficult to measure and link to productivity (see Brynjolfsson and Hitt, 2000, David, 1990). Robert Solow's oft-cited observation that one "can see the computer age everywhere except in the productivity statistics" is a good summary of this well known "Productivity Paradox."

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, where the 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. Moreover, the adoption of IT might coincide with other new management practices that are unobserved by the researcher (see Bartel et al., 2007).

A few papers address these issues focussing on specific applications of IT.[5] I follow this approach, analyzing a recently popularized IT innovation that is quickly spreading across police departments worldwide (Grossman et al., 2011). The innovation, called "predictive policing," collects and analyzes data on past criminal events to predict futures ones. Police patrols are given these predictions and are asked to reorganize their driving accordingly to increase their clearance or arrest rates and lower crime (Weisburd et al.,

[2]See the 1987, 2003, and 2013 Law Enforcement Management and Administrative Statistics (LEMAS).

[3]See, among others, Acemoglu et al. (2007), Autor et al. (1998), Berman et al. (1994), Black and Lynch (2001), Bloom et al. (2012), Bresnahan et al. (2002), Doms et al. (1997), Stiroh (2002).

[4]See Angrist and Lavy (2002), Athey and Stern (2002), Goolsbee and Guryan (2006), and, for police management, Garicano and Heaton (2010).

[5]See Athey and Stern (2002) and Hubbard (2003).

2003, see, for example,). Regardless of the growing interest and growing investments in predictive policing, several stakeholders have highlighted that very little is known about its effectiveness (Sengupta, 2013). The main empirical issues are the endogeneity of its use and the possible existence of displacement effects (criminals have an incentive to defy these predictions, and one way is to simply move away from locations that are predicted to see surges in crime).

In order to address these issues, I exploit *individual* offense level randomization in the availability of predictive policing. In particular, I use *micro-level data* on the universe of commercial robberies against businesses in Milan (Italy) over a two-and-a-half-year period,[6] and quasi-random assignment of predictive policing to *individual* criminal events to examine the empirical relationships between IT use and productivity of police patrolling, as measured by the likelihood that individual crimes are cleared by arrest.[7].

I develop *two alternative* difference-in-differences strategies. Both alternatives share the first difference: in Milan (as well as in all major Italian cities) *two separate* police forces are patrolling the streets, the *Polizia* and the *Carabinieri*, and only one developed a predictive policing strategy. Moreover, there is quasi-random assignment of investigations to these two separate police forces that is driven by very peculiar rotating mechanism (see Mastrobuoni, 2014). The city is divided into three sectors and approximately every 6 hours, when the shifts are changing, the two police forces are assigned to different sectors.

Even though the two forces share similar staffing and equipment (see Section 3.1) and have access to the same information (including the possibility of interviewing the victims), this difference would not be sufficient to identify a productivity change if, irrespective of

[6]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). 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.

[7]Several economic studies have used clearances as a measure of police performance (see, among others, Garicano and Heaton, 2010, Mas, 2006).

predictive policing, *Polizia* and *Carabineri* differed in their underlying productivity.

There are two *alternative* second differences I exploit to control for separate productivity levels. The first "second" difference is based on the very nature of predictive policing: the analysis of past criminal events. Any difference in clearance rates for the very first robbery of a sequence would be evidence of a differential productivity that is not based on IT. The second "second" difference is based on a procedural delay in producing the crime predictions. The *Polizia* requires time to collect and analyze the data. In order to optimize the victims' recollections officers wait about one day before interviewing the victims. This implies that predictions are not updated on the same day a robbery has taken place, generating a discontinuity in the availability of updated predictions.

While there is no evidence of a productivity differential between *Polizia* and *Carabineri* for the very first robbery of a sequence, subsequent robberies that fall in the *Polizia* sector as opposed to the *Carabineri* sector are 8 percentage points more likely to be solved (the overall clearance rate is 14 percent). Similar productivity differences emerge between robberies that happen before and after the predictions are updated, as long as the crime happened in an area surveilled by the *Polizia*. The results are robust to narrowing the sample to robberies that happen around the time the software is updated. Again, no differences emerge for the *Carabinieri* and between *Polizia* and *Carabineri* before the data update takes place.

I also provide evidence on the mechanisms that drives the productivity effect. Individual criminal behavior shows clear signs of predictability.[8] 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. Moreover, robbers tend to be very criminally active, which implies that at any given point in time the *Polizia*

[8] 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.

focuses on a small number of sequences.[9] I show that the instructions distributed to the police patrols highlight these patterns.

The large productivity boost in terms of clearances is expected to translate into more incapacitation and lower crime rates. Evidence based on auxiliary monthly municipality-level bank robbery rates provided not by the *Polizia* but by the Italian Banking Association shows that around the beginning of 2008, when predictive policing was first introduced, Milan robbery rates compared to rates in *any* other major Italian municipality experience a very sharp and abrupt reversal of a previously increasing trend. I conclude the analysis with a conservative cost benefit analysis where, even in the absence of deterrence effects, predictive policing appears to be very cost-effective.

As previously mentioned, this paper contributes to the literature on IT and productivity. A few studies have examined micro-level empirical relationships between IT use and productivity. Athey and Stern (2002) use a difference-in-differences setup to evaluate the effect of enhanced 911 emergency response systems–that link caller identification to a location database–on health outcomes. The IT adoption is shown to generate significant improvements in the health status of cardiac patients. Hubbard (2003) uses a conditional independence assumption to test whether trucks that use on board computers are more productive. Onboard computers (GPS, etc.) are shown to significantly increase the ability to predict the availability of trucks and therefore their capacity utilization. The paper also contributes to the growing literature on the mechanisms through which policing reduces crime, which I discuss in more detail in the next section. It also has implications about data collection for law enforcement agencies.

Section 3 discusses the identification strategy, Section 4 shows the results, and Section 5 shows evidence of persistence in criminal strategies. Policy implications as well as evidence about the aggregate reduction in robberies are presented in Section 6, while

[9] Sixty percent of matched offenders commit a new robbery within one week, 77 percent within two weeks, and 85 percent within one month. As a result, each month the average number of unique groups that are active and whose actions need to be predicted is around 13.

Section 7 concludes.

2 Predictive Policing

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). The data are often used to map crimes patterns and reorganize police patrolling.[10]

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 police patrolling in geographic areas (e.g., blocks or specific addresses) that show high levels of criminal activity. Hot-spots policing has gradually evolved from using data to simply identify high crime areas into a more advanced and dynamic information technology that uses higher frequency local crime rates to make predictions about future aggregate criminal activity.

Common predictive policing strategies use more advanced statistical techniques that are built on autoregressive models over time and space.[11] The most advanced ones predict the most likely type, location, and time of future crimes.[12]

Recently the National Institute of Justice (NIJ) has launched a demonstration initia-

[10]Garicano and Heaton (2010) study 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, which was developed by the New York Police Department's Police Commissioner William Bratton under Mayor Rudolph Giuliani's leadership.

[11]A few studies have evaluated hot-spots policing strategies, and most criminologists believe that focussed policing works (Braga, 2001, Cohen and Ludwig, 2003, Sherman and Weisburd, 1995, Weisburd and Green, 1995). Levitt (2004) is more skeptical about the decline in crime that occurred during the 1990s that can be attributed to Compstat, and while there is currently little evidence that hot-spots policing simply displaces crime to nearby locations, 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).

[12]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).

tive 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).[13]

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). The experiment seemed to reduce crime, though the absence of a control group and the possibility that crime was merely displaced make it difficult to draw any definite conclusions.[14]

The issue in most of these studies is that they either lack a proper comparison group. Criminals might move from treated to control regions contaminating the experimental design. Reducing contamination by choosing larger regions would introduce additional heterogeneity between treated and control areas. 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

[13] 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.

[14] 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.

over time.[15]

Combining the Italian institutional background with the micro-level empirical strategy overcomes these issues, allowing for an evaluation of the productivity effect of IT use that is not prone to such displacement. The micro-level data allows me to shed light on 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 of criminals, commuting patterns, etc.), the characteristics and activity of victims (e.g., victims' location, their vulnerability, etc.), and the interplay between criminals and victims (e.g., gang shootings, victims' precautions, etc.).[16]

2.1 The Milan Police Predictive Policing and the Data

In 2008 the Milan Police Department (*Polizia*) started implementing a software called "Keycrime" that collects and analyzes micro-level data on all commercial robberies that take place in the municipality of Milan (*Comune di Milano*).[17] The predictive policing software is used to input and analyze large sets of individual characteristics of robbers and individual criminal strategies (*modus operandi*) collected from closed-circuit security cameras and victim reports in order to: i) identify robberies that share at least one offender or one vehicle (called a "sequence");[18] and ii) predict when and where the offenders are going to strike next.

[15]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 (2014) 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.

[16]See Clarke (1997) for a discussion about situations that enhance criminal opportunities.

[17]Commercial robberies, which are crimes of violence against businesses motivated by theft, are quite prevalent in Italy. Bank robberies, which comprise about 10 percent of all commercial robberies, are more prevalent in Italy than in the rest of Europe altogether (see Mastrobuoni, 2011).

[18]The linkages across robberies are constructed irrespective of whether an arrest is made (see Section 2.1).

The *Polizia* uses the information produced by the predictive policing software for two purposes. The first is to produce detailed instructions for police patrols. The second is to assist the prosecutors once the perpetrators have been arrested and are put on trial.[19]

2.1.1 The IT Innovation

After a robbery takes place the predictive policing team collects and later examines around 11,000 bits of information about the crime (time, date, location, type of business, type of crime, etc.), about the observed perpetrators (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 by the perpetrators (type, maker, model, license-plate, etc.).

The *Polizia* first gathers data about the event collecting the official *Polizia* or *Carabinieri* reports, and later interviews victims , and collects surveillance camera footage. About 80 percent of businesses have closed-circuit security cameras (CCTV).[20] Between January 2008 and June 2011 the *Polizia* has recorded around 2000 robberies, at a rate of 1.5 robberies per day.

The *Polizia* collects this information for the universe of reported commercial robberies that take place in Milan. Given the monetary and non-monetary incentives to report these crimes (many businesses are insured and understand that future patrolling strategies may depend on their reporting behavior), reporting rates among commercial businesses are believed to be close to 100 percent.[21]

[19] 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 later represents our control group, made the arrest.

[20] According to the *Polizia* all banks, postal offices, pharmacies, and jewelers have at least one CCTV camera.

[21] According to the *Polizia* only in one instance did a robber confess to a robbery that had not been reported.

The interviews of victims, which represent the core of the information collected, happen over the phone the day after the robbery has taken place. The purpose of the delay is to reduce the victim's post-traumatic stress disorder and improve their recollections.[22] After collecting and inputting the data into "Keycrime," the software is used to ease the operators' job of matching robbers or group of robbers over time. Figure 9 shows that the software allows the operator to compare on one screen the characteristics, including the photographic evidence, of different robberies.[23]

Once links are established (later I discuss the possibility that the links might be misclassified), the data are used to highlight and to predict criminal strategies. The predictions are based on a mix of statistical, psychological/criminological models.[24]

The potential future targets are then communicated to police patrols (some are indicated in Figure 11 with a small blue square), 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 10. 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. 10 has presumably robbed 22 business, which is why such evidence is particularly rich.

At the beginning of 2010 the prosecutors asked the *Polizia* to share the reports with the *Carabinieri*, which might have pushed the *Carabinieri* to develop similar policing strategies.

[22]Later I exploit such delays to setup the second difference-in-differences strategy.

[23]While I do not have access to the proprietary algorithm that predicts criminal behavior, I have been told that the current pattern recognition softwares are not yet capable of automatizing the matching of photographic evidence. Moreover, when the evidence is missing the operator can still use the mutual appearance of peculiar and rare physical appearances to establish these links.

[24]The algorithm is proprietary and I do not have access to it, but some simple theories are tested in Section 5.

2.1.2 The Data

I have been given access to the data collected through "Keycrime" between January 2008 and June 2011, with great detail on the *modus operandi* of the robberies (location, time, loot, arrest, number of offenders, weapons, type of business, etc.).[25]

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 were 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, that is before the *Polizia* started sharing their predictions with the *Carabinieri.* 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 vehicle were seen in two different instances. The variable "Number of the sequence" $n = 1, ..., N_i$ counts the number of crimes that have been linked to a serial group of offenders i.[26] The criminal group with the largest number of offences organized 84 robberies.

The *Polizia* defines a given robbery to be cleared if an arrest is made before the same group of robbers re-offends.[27] A sequence i is believed to be solved when all observed robbers have been arrested.

More than half of the robberies (1,221 robberies out of 2,164) belong to a sequence where at least one arrest has been made. Of these, 981 (80 percent) belong to a sequence that has presumably been fully cleared.[28]

[25] Since the predictive policing software, which is used to collect the data, started to be fully implemented in 2008, there are no data on clearances by the *Polizia* and the *Carabinieri* available before that year.

[26] For the serial crimes that started in 2007 and continued in 2008 I have the number of robberies performed in 2007, which I added to the "Number of the sequence."

[27] I do not have complete information on the exact date of arrest, but according to the *Polizia* considerably more than half of all arrests happen *in flagrante*, meaning when a robbery is taking place, or when the robbers are fleeing.

[28] Though, in principle it would still be possible for the sequence to proceed if new perpetrators were using the same vehicles used by the arrested ones.

Table 1 shows that the individual clearance rate of robberies is 14.9 percent, which leads to 45 percent of the sequences being fully cleared by June 30, 2011, the day the data were extracted. The *Polizia* variable indicates whether the Milan Police Department handled that particular robbery and the next Section is going to describe how this assignment of investigations to the *Polizia* and the *Carabinieri* works. Since the city is divided into 3 parts and the police is responsible for 2 of these, the fraction of robberies handled by the police is slightly larger than expected (73 percent against 67 percent).

According to the victims' reports the robbers appear to be on average 26 years old. The average haul is around €2,000, or $2,200. 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 the quasi-experiment.

3 Experimental Design

Part of a crime reducing effect of focussed policing strategies is likely due to incapacitation: preventing subsequent crimes by captured criminals. Incapacitation lowers the crime by the counterfactual number of offenses that detained criminals would commit had they not been arrested. This reduction is likely to generate diffused benefits over time, making it hard to infer from simple pre-post differences in crime rates the effect of predictive policing, or as a matter of fact any focussed policing. Even if criminals are unaware of the IT innovation, as long as they are sufficiently mobile, treatment effects based on contemporaneous differences across locations would also be biased towards zero.(see Cook, 1979, Nagin et al., 2015)

An alternative identification strategy is to randomly assign predictive policing at the police patrol level, and to measure the effect on clearances rather than on crime rates. If the assignment is random any deterrence would be diffused, and differences in clearance rates would be ascribable to productivity improvements. 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 experimental design.[29]

Even with such a design, it would still be necessary to randomly assign the crimes that need to be investigated. Otherwise, an increased productivity (higher clearance rate) might just be driven by treated patrols cherry-picking the more predictable and potentially poorly organized crimes, overstating the effectiveness of predictive policing. Assigning patrolling areas to treatment and control patrols is probably the most natural way to avoid such cherry-picking. 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 earlier.[30]

The next section describes how the IT operative rules of predictive policing and the existence of two separate police forces are combined to design a quasi-experimental identification strategy of productivity effects.

[29]The experimental assumption is known as the Stable Unit Treatment Value Assumption (SUTVA).

[30]Another 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.

3.1 First Difference: Two Police Forces

For historical reasons, Italy has two separate police forces:[31] 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.[32] When investigating commercial robberies the two forces differ in the availability of predictive policing, which is not part of a wider set of innovations, but rather a product of a number-crunching police officer.

In all major cities the two forces operate side by side, without communicating with each other.[33] Moreover, the above-mentioned cherry-picking is avoided by the fact that cities are divided into three different areas (two falling under the responsibility of the *Polizia* and the third of the *Carabinieri*), and each force is solely responsible for keeping law and order in the assigned area. On its own, even such division into areas would not provide random variation in crimes, because forces could be assigned to the zones according to their productivity, or criminals could 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 patrols to the three areas rotates approximately every 6 hours, counterclockwise (at 12am, 7am, 1pm, and 7pm). Given that there are two forces, three areas, and four 6-hour shifts within a given day, patrols belonging to one police force 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

[31]See Mastrobuoni (2014) for a discussion about the two forces.

[32]The only difference between the two forces is that the *Polizia* operates exclusively in metropolitan areas, while the *Carabinieri* operate on the entire Italian territory. This difference is not going to influence this analysis as I am going to compare forces that operate within the boundaries of the city of Milan. While the *Carabinieri* might have an advantage when investigating criminal groups that operate both inside and outside of city, according to the *Polizia* the mobility of criminals in and out of the city is limited.

[33]The communication aspect started changing in January 2010, when prosecutors asked the *Polizia* to share their predictions with the *Carabinieri*. In January 2010, the *Polizia* started sending information to the *Carabinieri*, and by the end of the year they had shared 33 classified reports. Section 4.1 shows that staring in 2010 the productivity of the *Carabineri* does indeed converge to the one of the *Polizia*.

coverage of police patrols. Figure 1 shows the distribution of robberies in Milan based on the day triplet, where the robberies that are under the responsibility of the *Carabinieri* are shown with a black square and the ones that are under the responsibility of the *Polizia* 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 *Carabinieri* patrols cover the northwestern part of the city while the *Polizia* patrols cover the rest. In group 2 day/time combinations the *Carabinieri* cover the northeastern part and in group 3 the southern one.[34]

A first difference between the ideal experiment and the actual one is that the assignment of patrolling areas to police forces, and therefore to predictive policing, follows a predetermined pattern. This implies that criminals could potentially target areas that are not patrolled with the aid of predictive policing. Such an endogenous response would lead to quantitative and qualitative sorting, with more crimes as well as more professionally organized crimes falling in the "untreated" areas. The endogenous response of criminals will be tested in Section 4.4.

A second difference is that treatment (*Polizia*) and control (*Carabinieri*) patrols might, irrespective of predictive policing, differ in their productivity. 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 Romeo 159, 2.4 JTDM 20v with 200 horsepower, is their standard service car, see Figure 12), they are almost identical in size. According to law, nationwide there are 57,336 police officers and 48,050 *Carabinieri* 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 *Carabinieri*), and the numbers

[34] The few outliers are driven by robberies that have been assigned to i) police or *Carabinieri* 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, 2014). There are so few outliers that none of the results depends on them.

of top-rank officials are similar as well. Yet, a credible identification strategy would have to difference out any underlying productivity differences, which is the objective of the second difference.

3.2 Second Difference

3.2.1 First vs. Subsequent Robberies

I use two alternative ways to get rid of underlying productivity differences. The first strategy separates robberies between first and subsequent events of a sequence. The reason for separating first and subsequent offenses is that one would not expect predictive policing to work without having previously gathered the data. The probability of clearing a robbery might differ between first and subsequent robberies for other reasons too. There is likely to be a strong selection if the most inept robbers are immediately caught. And the ones that are not might also learn with experience. But there is no reason why these differences should differ across the two police forces, unless smarter robbers choose the less productive police force (which is tested in Section 4.4).

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 within a sequence.[35]

In order to control for confounders X, I model clearances using a linear probability model, where the dummy variable is equal to one when the n-th robbery within a sequence i is cleared before the next robbery takes place:

$$Cleared_{i,n} = \alpha + \delta_k Polizia_{i,n} + \gamma' X_{i,n} + \epsilon_{i,n} \,; k = 1(n > 1). \quad (1)$$

[35] In principle, when analyzing photographic evidence the police might recognize individuals. But this would happen irrespective of the force that is operating on the ground and, thus, would not be able to explain differences in productivity.

The coefficient δ_k on *Polizia Intervention* measures the simple difference in clearance rates between the *Polizia* and the *Carabinieri*. Combining the simple differences in a difference-in-difference setup:

$$Cleared_{i,n} = \alpha + \delta_0 Polizia_{i,n} + \delta_1 1(n > 1) + \delta_2 Polizia_{i,n} \times 1(n > 1) + \gamma' X_{i,n} + \epsilon_{i,n}. \quad (2)$$

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. 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 is simply

$$Cleared_{i,n} = \alpha + \delta_0 Polizia_{i,n} + \delta_1 n + \delta_2 Polizia_{i,n} \times n + \epsilon_{i,n}. \quad (3)$$

There are two potential issues with the identification strategy based on n, the number of the sequence. The first is that the *Polizia* officers might be getting better and better as n grows simply because they collect information across the entire city at all times while the *Carabinieri* officers restrict their work to the assigned areas at the assigned times. In other words, the improvement in clearances might also be related to the data collection itself. While nothing prevents the *Carabinieri* from collecting the same kind of data, the next identification strategy, which is based on the time the software is updated, does not rely on a comparison between the *Polizia* and *Carabinieri*.

A second issue is that the number of the sequence (n) might be misclassified, potentially biasing the results. Errors when linking different robberies over time, and, therefore, errors in the measurement of first and subsequent robberies would lower the accuracy of the predictions and the estimated efficacy of predictive policing; unless such errors are systematically and differentially linked to clearances. The next identification strategy restricts the attention to subsequent robberies and does rely on differences that are based on n.

3.2.2 Same Day vs. Different Day Robberies

To reduce victims' post-traumatic stress and, consequently, avoid any recall bias about the robbery, the *Polizia* waits until the following day before interviewing the victim. This means that when robbers perform two robberies in one day, patrols are not going to have an updated version of the prediction for the second robbery until next day. This identification strategy *does not* exploit differences between the first and subsequent robberies but only differences *within* subsequent robberies based on their timing.[36]

The estimated equations resemble Equations 1 and 2, with the binary variable

$$1(\text{different day robbery}_{i,n}) = \begin{cases} 1 \text{ if the robbery does not happen the same day;} \\ 0 \text{ otherwise.} \end{cases}$$

replacing $1(n > 1)$. Trimming the window around the time the software is updated I can compare, for both police forces, the likelihood of clearing a robbery just before and after the update.

4 Results

4.1 First vs. Subsequent Robberies

Table 2 shows the clearance rates by year and by police forces, separating robberies between first and subsequent events of a sequence. Overall there are few differences in clearance rates between *Polizia* and *Carabinieri* for first events in a sequence. For subsequent events, instead, clearance rates are much higher for the *Polizia*. Only after

[36] If clearance rates for the *Polizia* were upwards biased, this would be true irrespective of the timing of the second robbery; and any relative increase in the probability of clearing a case would still be unbiased.

2010, when the *Polizia* started sharing its knowledge with the *Carabinieri* do clearance rates converge again. Moreover, the curly brackets shows the fraction of robberies that fall in each area, and there is no evidence that robbers target areas that are patrolled by the *Carabinieri* more frequently; even when one focusses on subsequent ones, where the productivity differences are large, it does not seem that criminals are trying to operate in areas that are patrolled by the *Carabinieri*.

That the differences for those two years are significantly different from zero can be seen in the simple difference-in-differences Table 3, where, in addition to Table 2, subsequent robberies are separated by their number of the sequence. Since, naturally, the sample size shrinks with the number of the sequence, the differences when the number is grater or equal to 4 are lumped together. The only difference that is not significantly different from zero is for the very first robbery of a sequence ($n = 1$). All the other differences are positive and significant. There is clear evidence that between the first and the second robbery in the absence of predictive policing clearance rates drop considerably, which is consistent with both, selection and learning.

The estimated coefficients of Equation 1 and 2 are shown in Table 4. Columns 1 and 2 restrict the analysis to first robberies ($n = 1$), while columns 3 and 4 to subsequent ones ($n > 1$). The difference in clearance rates is close to zero among first robberies and is equal to 10 percentage points (significant at the one percent level with standard errors clustered by criminal group) for subsequent ones.

Consistent with the quasi-random assignment of police forces to crimes, controlling for additional regressors listed at the bottom of the table (columns 2 and 4) leaves the coefficients almost unchanged. Relative to the *Carabinieri* these results mean that the *Polizia* officers are almost 3 times more likely to solve subsequent robberies compared to the *Carabinieri* 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.

In Column 5, when computing the difference-in-differences between *Polizia* and *Carabinieri* for first and subsequent robberies, the effect of predictive policing is equal to almost 8 percentage points. In Column 6, allowing the effects of predictive policing to depend on the number of robberies increases the fit of the model as well as the precision of the estimates. When there is a *Polizia Intervention* the likelihood of clearing a case increases by 0.9 percentage points (more than 10 percent) for each additional robbery (*Number of the sequence*) the predictive policing software can analyze. The average number of subsequent robberies is 8, so the average predicted difference is about 0.07, which is close to the difference estimated in Column 5.

It is also important to notice that for the *Carabinieri* the coefficients on "Subsequent robberies" and on the "Number of the sequence" are negative, indicating that, either due to selection, or due to learning, successful robbers become more and more difficult to arrest. Predictive policing counteracts these forces.

4.2 Same Day vs. Different Day Robberies

Next, I exploit the procedural lags in collecting the data, a strategy that does not rest on the correct measurement of the number of the sequence. The left and right panel of Figure 2 show the clearance rate for the *Polizia* and the *Carabinieri* depending on whether the second robbery happens within the same day (a lag of 0 days), a few days later (1 to 5), or 6 and more days later. Due to the small sample size beginning with lag 2 I smooth the series using a moving average of order one.[37] The smoothing is particularly important for the *Carabinieri* where the sample size is about 30 percent of what it is for the *Polizia*.[38]

The squares indicate the average clearance rates, the vertical bars the corresponding

[37]The smoothed clearance rate is equal to c_t for $t \in \{0, 1\}$ and $\tilde{c}_t = \frac{c_t + c_{t-1}}{2}$ for time $t \geq 2$.
[38]In order to maximize sample size here I use all the data (2008-2011).

95 percent confidence intervals. The horizontal line corresponds to the average when lumping together all "subsequent day" robberies that do not happen on the same day. For the *Polizia* clearances jump from less than 5 percent to more than 15 percent when the robberies happen one day later rather than on the same day (there are 88 robberies in the first category and 125 in the second). For the *Carabinieri* clearance rates are more noisy, as the sample size is smaller, but there is no evidence of an increase in clearances once a day has passed from the previous robbery.

The regression coefficients that measure differences based on the updates of the software are in Table 5. Columns 1 and 2 show that the *Polizia*'s productivity is considerably larger when robbers do not perform their subsequent robbery within the same day. For the *Carabinieri* no such difference emerges (columns 3 and 4). Conversely, within the same day the two police forces have very similar clearance rates. The difference-in-differences are shown in Columns 5 and 6 and are slightly larger than the difference-in-differences based on subsequent vs. first robbery identification strategy.

Consistently with the evidence shown in Figure 2, when I restrict the comparison to robberies that happen within a few days from the previous robbery the results stay the same (see Table 6).[39] The *Carabinieri*'s differences in clearance rates between same day and different day robberies are very close to zero.

Since data updates generally happen around 10am, one can use a regression discontinuity design around such time. The minimum time since update is -31 hours, which corresponds to a robbery that took place at 3am in the morning and was updated 31 hours later. Table 7 presents such estimates: columns 1 to 4 for the *Polizia* and columns 5 to 8 for the *Carabinieri*. The first two columns in each panel cap the time since the data update at 120 hours (5 days) and control in addition for the running variable,[40] while the

[39] The results in Columns 1 and 5 are not exactly the same as those in Table 5 because to increase power I am using all years, again potentially biasing the coefficients for the *Carabinieri* upwards.

[40] AIC and BIC model selection criteria reject the existence of separate slopes on both sides of the threshold.

following two columns cap the time since the data update at 32 hours, without controlling for the running variable. Columns 2, 4, 6, and 8 exclude the robberies that fall within 3 hours of the discontinuity, where the assignment to the data update status is more uncertain (giving rise to the "donut" regression discontinuity). In order to control for time differences on the left and on the right of the discontinuity all regressions include four "police shift" fixed effects (6 hour intervals). The results are very much in line with the daily differences. *Polizia* and *Carabinieri* have the same productivity before *Keycrime* is updated and the new instructions are distributed to the patrols. Once the 10am threshold is hit clearances increase by about 10 percentage points, but only for the *Polizia*. The donut RD tends to increase the discontinuities, which is in line with the fuzziness of the RD around the 10am threshold.

Such a rapid response of predictive policing brings credit to the hypothesis that the effects are driven by updated police reports leading to better focussed patrolling. If, instead, the effects were driven by investigations being aided by a more thorough data gathering the timing of the effects would most likely be less immediate.

Summing up. all difference-in-differences and the regression discontinuity give very similar results: no productivity differences when predictive policing is either not available or has not yet been updated, and large productivity differences once predictive policing is fully operational. This suggests that misclassification in the number of the sequence is not biasing the first set of results.

The next Section performs a series of robustness checks, which for brevity and for reasons of statistical power are based on the first set of results.

4.3 Robustness

This section addresses a few common robustness checks, including functional form assumptions, outliers, spillovers and heterogeneity of the effects. Column 1 of Table 8 indicates that marginal effects of a probit model are in line with the linear probability estimates.[41]

To address the heterogeneity of results based on the availability of CCTV cameras, in Column 2 I only include robberies where it is known with certainty that current or past victimized commercial robberies have some CCTV cameras installed.[42] The results are indeed a little larger, suggesting that the availability of CCTV cameras might benefit the predictions. Column 3 shows that the predictive policing effects are only slightly lower when focussing on robberies where the loot is above average. This may indicate that robberies with lower loots are easier to predict. Excluding the single most victimized category, pharmacies, does not alter the results (Column 4), indicating that the most victimized businesses are not driving the results. Finally, in the last column there is no evidence of biases due to spillover effects. When focussing on just the very first robbery of the day a police force has to deal with, the results are unchanged.

4.4 Testing for an Endogenous Response of Criminals

It is unlikely that robbers would know about the exact timing when the data are collected and the software is updated, suggesting that the identification strategy based on the timing of the update is not prone to selection bias.

But in the identification strategy that compares first vs. subsequent robberies capable and experienced robbers might try to target victims when their business falls in the *Carabinieri* area. Starting with a balance test, Table 9 compares all the observable characteristics of the robbers and of the robberies depending on whether the *Polizia* or the *Carabinieri* were 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

[41] The same is true for the differences based on the software updates.

[42] I was not given any photographic evidence but was told that all banks, postal offices, pharmacies, and jewelleries have CCTV systems that are constantly running.

robbers (see Mastrobuoni, 2011) does not show any differences. One cannot reject the hypothesis that not just the mean of the loot but its entire distribution is the same for the two forces (see the Online Appendix Figure 13). There are a few variables for which Table 9 measures small but significant differences between the two forces: year, day of the week, pharmacies and other businesses. Controlling for these small differences did not matter.

Finally, if robbers knew about these differences there should be relatively more robberies that fall under the responsibility of the *Carabinieri* than of the *Polizia*, especially past the first robbery (when predictive policing might potentially aid the investigation).[43] 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 *Carabinieri*.

Despite the lack of an excess mass in the number of subsequent robberies that fall into the *Carabinieri* area, and the balance in the loot between the two forces, my final specification is again going to deal with the possibility that more able robbers sort into the *Carabinieri* zones. If such a selection was driving the results, sorting in the previous robberies would also be predictive of differences in the likelihood of a clearance as the unobserved ability of robbers that drives the selection would be a persistent trait. Table 10 shows that there is no evidence of this. Regressing the *Cleared robbery* dummy on the dummy whether the *Polizia* was patrolling the area as well as its first two lags reveals that only the last treatment matters. Moreover, Column 3 shows that the interaction between the current *Polizia* dummy and the past one has no impact on clearing the case. Moreover, a related implication is that the information collected by the *Polizia* officers when the *Carabinieri* officers covered the previous robbery is not inferior to the one they would have collected.

[43] The summary statistics Table 1 already showed that robberies do not seem to disproportionately target the *Carabinieri* patrolling areas.

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 research 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. The micro-level data allows me to analyze these mechanisms.

5 Evidence about the Mechanism

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 of individual robbers or group of robbers using all years of data and all the variables that I have been given that could potentially be exploited by a predictive policing software.

One way to show persistence in the choice of the location of a robbery is to plot these for each group of robbers. Figure 3 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 geographic clustering, most teams of robbers restrict their activities to certain neighborhoods. 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 one. If there was no persistence such likelihood would equal the marginal distribution of business types.

Figure 4 shows that the marginal distributions are 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 5 shows that a very similar pattern emerges when one classifies the time at which a robbery is committed 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 late afternoon.

Finally, Figure 6 shows that there is some persistence in the chosen day of the week, but only Sunday and Monday, and to a lesser degree on Friday and Saturday, possibly because of working schedules. 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 7 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.

When continuous variables (X) are used to measure (negative) persistence I compute their mean absolute deviation from the mean using only past and present data ($t^{-1}\sum_{\tau\leq t}|X_{i,\tau} - \overline{X}_{i,t}|$, 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 11 shows evidence that robbers who often select their modal hour of the day tend to select their modal type of business.

The 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 12 I regress each of these factors on a measure of experience (the *Number of the sequence*), 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, as no causal claims can be made. Robbers who have performed a robbery whose loot tuned out to be larger than average for that type of business are 3 percentage points more likely to choose 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 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 drops by 27 percent, the one about distance by 28 percent, and the one about time between the robberies by 20 percent.

6 Aggregate Evidence and Policy Implications

As mentioned earlier, clearing a robbery means that at least one robber is arrested.[44] Based on data collected by the *Polizia* close to 100 perpetrators were arrested between 2008 and 2009. Of these only one perpetrator was acquitted, while the rest received a total of 420 years of prison time (about 4 years per prisoner).[45]

After their first robbery about 30 percent of robbers are linked to a second robbery, and after that almost all are observed reiterating their crime until an arrest takes place (see Mastrobuoni, 2014, for evidence on this "life" table of robberies). For this reason differences in clearance rates lead to differences in the expected number of robberies criminal groups are able to organize before ending up in prison. Since the *Polizia* and the *Carabinieri* 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.

The two difference-in-difference estimates were 7.8 and 13.3 percent. Using an average of 10 percent as an estimate and setting the counterfactual clearance rate at 5.6 percent for subsequent robberies without predictive policing (this is the clearance rate for the *Carabinieri* in 2008 and 2009), the expected number of robberies each group of recurrent criminals commits drops from about 17.8 to about 6.4, a 2.8 to 1 ratio.[46]

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

[45] Four criminals were given alternative sanctions to prison time.

[46] When the criminal attempts are frequent and persistent, as it happens to be the case for Milan, the expected number of robberies is approximately equal to $\sum_{\tau=0}^{\infty}(1-c)^{\tau} = 1/c$, where c is the clearance

Since there are about 255 successful first time robbers a year and about one-third re-offend, the reduction of 11 robberies per sequence leads to a total reduction of 935 robberies (in the long run deterrence might lead to even larger reductions). Multiplying such number by the overall 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$ 2.8 million.[47]

The final empirical exercise will be to test whether these predictions square with evidence based on aggregate crime statistics that are not collected by the *Polizia*.[48] The Italian Banking Association has been collecting data on the monthly number of bank robberies for many years (bank robberies represent about 10 percent of all commercial robberies). I have been given access to the monthly series of bank robberies for all major municipalities (*capoluoghi di provincia*) between 2004 and 2015.[49] Figure 8 shows the time series of bank robbery rates per 100,000 inhabitants (based on 2006 population estimates) in Milan (left panel) and in Milan compared to the nine largest municipalities (right panel). The evidence shows an upward trend in bank robberies that reverses around the time Keycrime became operational (early 2008). Attributing all these changes to Keycrime based on just this figure would be subject to the same criticism mentioned in Section 2, e.g. mean reversion, but in conjunction with the micro-level evidence based on clearance rates it does paint a picture of a very effective policing strategy.

The changes are very large. During the time Keycrime has been evaluated for this study (2008 to 2011) robbery rates fell from about 1.4 to about 0.5, a 2.8 to 1 ratio, which is not very far from the 2.5 to 1 ratio predicted for recurrent criminals based on clearance rate.

[47]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 (Cornaglia et al. (2014) find evidence of such indirect costs.)

[48]The Italian statistical office (ISTAT) started releasing municipality-level yearly crime data only in 2010, and these data are also based on data provided by the *Polizia* and the *Carabinieri*.

[49]During the 11-year period 10 out of 107 major municipalities did not have a single bank robbery and are therefore excluded from the analysis.

rates. Any endogenous response of criminals as well as any mean-reversion would make it hard to compare the two estimates, and yet it is comforting to see that they are in the same ballpark.

Plotting the difference between the rates in Milan and the rates in nine separate cities does not alter these results, implying that the size of the trend reversion is unique to Milan. Notice that a gradual trend reversion is exactly what incapacitation would be generating, as more and more potential offenders are prevented from committing a sequence of bank robberies. Starting in 2012, probably as Milan transits to the equilibrium with improved clearance rates, the reduction in robbery rates appears to slow down.

These results do not change when using all major municipalities as a comparison group. Table 13 shows difference-in-difference estimates in levels and in changes over time of bank robbery rates when using 97 major municipalities. The evidence of a significant trend reversion for Milan around 2008, even controlling for time and municipality fixed effects, is quite strong. The specification in changes, which according to the discussion of Fig. 8 appears to be the correct one, implies a 3.7 percentage point reduction in the bank robbery rate, which compared to the peak rate of 1.5 bank robberies per 100,000 inhabitants, corresponds to about 2.5 percent per month.

Going back to the cost-benefit analysis, one needs to take into account the increased cost from incarcerating arrested criminals and the cost of investing in the IT. Since all robbers eventually end up in prison, or in other words, since $(1 - c_t)^\tau$ converges to 0 reasonably quickly for clearance rates that are close to 10 percent, predictive policing merely anticipates the timing of incarceration. Since most re-offending happens within a few weeks, predictive policing tends to anticipate arrests by a few months, at most a few years. The average time between robberies is 15 days and the number of prevented robberies is 11, so the average time to arrest drops by about 6 months. With a reduction of 6 months, an interest rate of five percent (an upper bound of the yields on the

Italian government bonds), and an average yearly cost per incarceration of €50,000 (see Barbarino and Mastrobuoni, 2014), the cost of moving forward the incarceration expenses would be at most 2.5 percent of €1.25 million, or €31,250.[50]

The labor cost of the three fulltime police officers who collect the data and predict the crimes is below €100,000 a year. The investments in capital (an office, computers, monitors, etc) are hardly 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. Overall, the cost of introducing predictive policing appears to be an order of magnitude lower than the benefit.

7 Conclusions

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. Once the data to be analyzed become available, either because a history of criminal events develops or because the officers have the time to process the new information, the differences in productivity are striking. The micro-level information shows that these productivity differentials are consistent with the criminals' observed persistence in criminal behavior. Over time recurrent robbers tend to target similar businesses, around the same neighborhood, and at the same time of the day; together with the fact that robbers recidivate at a very high frequency (60 percent are "back in business" within one week), makes robbers predictable.

[50] Victimizations, as well as incarcerations, generate pain and suffering which I do not attempt to quantify, as both are extremely hard to measure.

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.

The experimental design allowed me to estimate the effect of predictive policing on the likelihood that a robbery is solved and a perpetrator is arrested. An open question is whether over time, as the productivity of policing is perceived to go up, predictive policing either deters crime altogether, convinces criminals to switch to other crimes, or displaces crime from Milan to other cities. Another open question is whether the *Polizia* is using the best possible prediction algorithm and whether there are ways to improve such predictions using more or even less detail about the robberies. Predictions based on such detail, where labor input is relatively high, may become unfeasible when the number of crimes grows larger. Whether there is a threshold level of complexity where it is better to aggregate the predictions is another open question.

In conclusion, this papers adds to the limited micro-level evidence on the positive productivity effects of IT investments (Athey and Stern, 2002, Hubbard, 2003). It is also the first quasi-experimental evaluation of predictive policing. These IT investments can be highly effective in improving the productivity of the police officers in their role of apprehension agents (Nagin et al., 2015).

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.

Gediminas Adomavicius and Alexander Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *Knowledge and Data Engineering, IEEE Transactions on*, 17(6):734–749, 2005.

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.

Nicholas Bloom, Raffaella Sadun, and John Van Reenen. Americans do it better: Us multinationals and the productivity miracle. *American Economic Review*, 102(1):167–201, 2012.

Richard J Bolton and David J Hand. Statistical fraud detection: A review. *Statistical science*, pages 235–249, 2002.

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.

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

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

Philip J Cook. The clearance rate as a measure of criminal justice system effectiveness. *Journal of Public Economics*, 11(1):135–142, 1979.

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

Francesca Cornaglia, Andrew Leigh, and Naomi Feldman. Crime and mental wellbeing. *Journal of Human Resources*, 49(1):110–140, 2014.

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.

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.

Thomas N. Hubbard. Information, decisions, and productivity: On-board computers and capacity utilization in trucking. *American Economic Review*, 93(4):1328–1353, 2003.

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. 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.

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

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. Police and Clearance Rates: Evidence from Recurrent Redeployment Within a City. Technical report, 2014. mimeo.

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.

Daniel S Nagin, Robert M Solow, and Cynthia Lum. Deterrence, criminal opportunities, and police. *Criminology*, 53(1):74–100, 2015.

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.

Somini Sengupta. In hot pursuit of numbers to ward off crime. *The New York Times*, June 19 2013.

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.

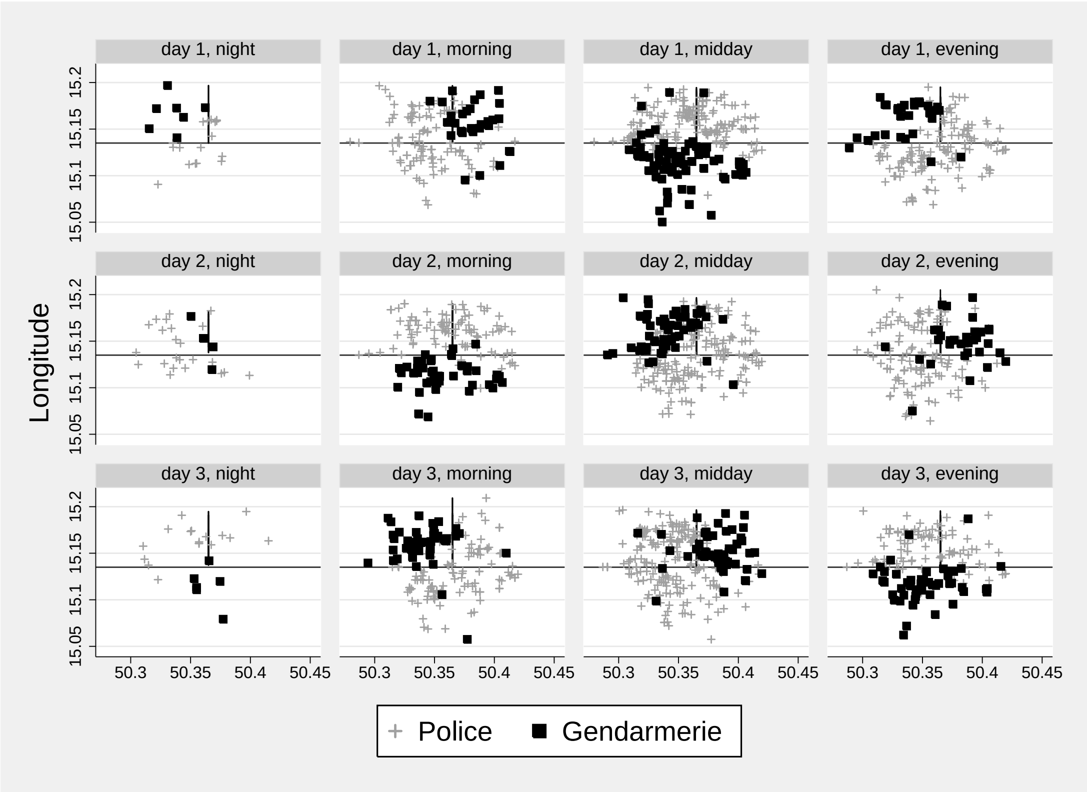

Figure 1: Geographic Distribution of Robberies by Group

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

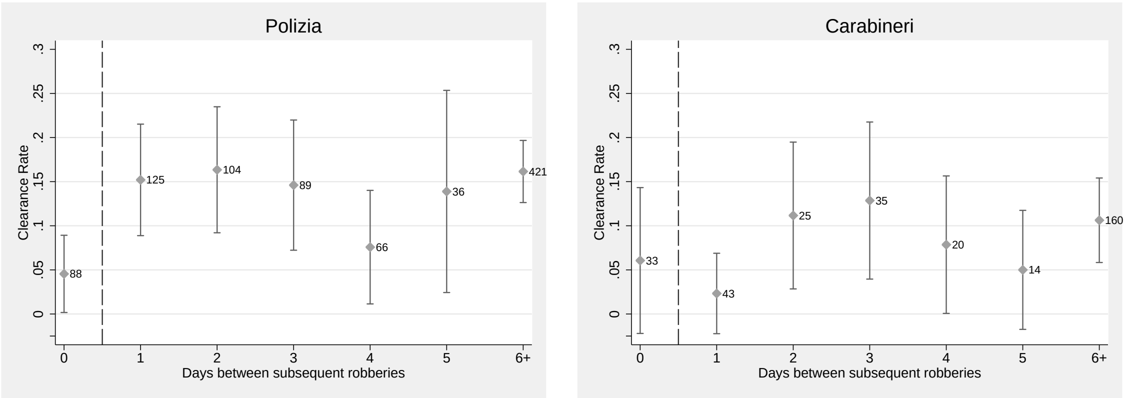

Figure 2: Clearance Rates by the Number of Days Since Previous Robbery

Notes: The numbers next to the average clearance rate indicate the number of observations. From lag 2 on ($t = 2$) the estimates are based on a simple moving averages of order one ($\hat{r}_t = \frac{\overline{r}_t + \overline{r}_{t-1}}{2}$). The vertical lines show the 95 percent confidence intervals.

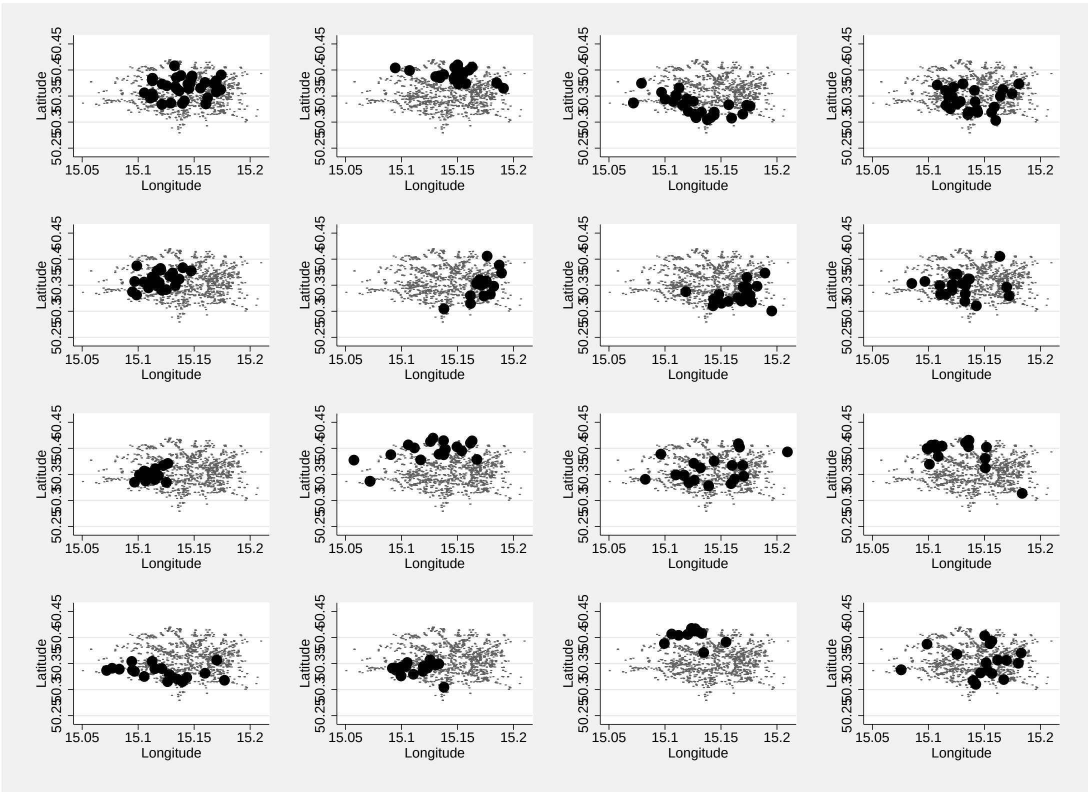

Figure 3: Geographic Distribution of Robberies by Criminal Group

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.

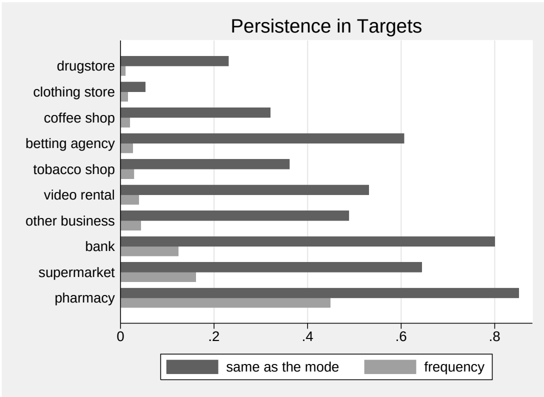

Figure 4: Persistence in Targets

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.

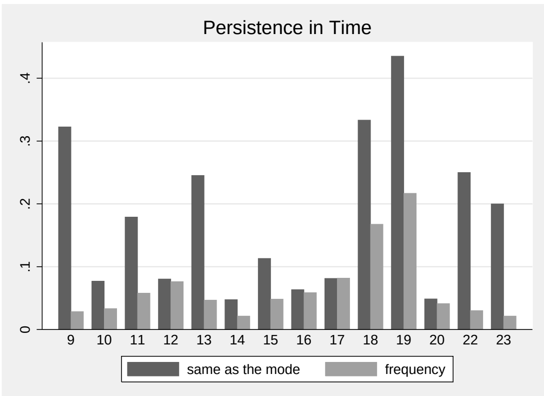

Figure 5: Persistence in Time

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.

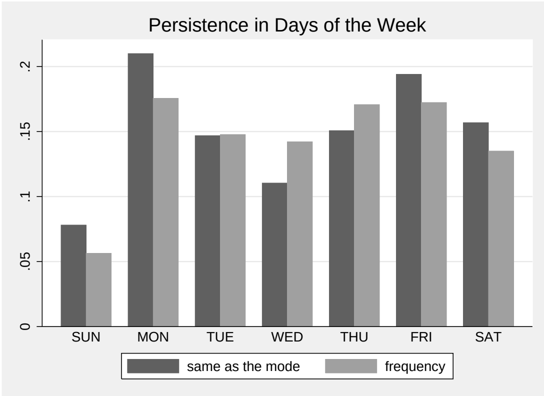

Figure 6: Persistence in Days of the Week

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.

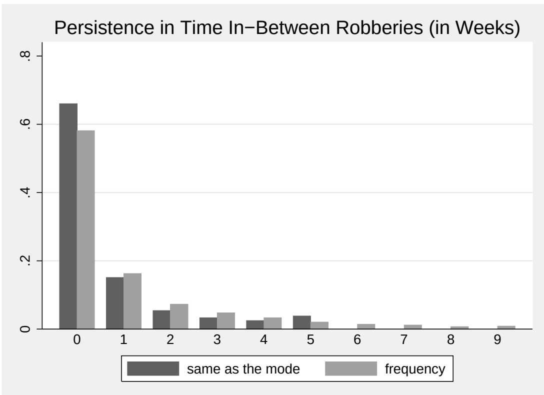

Figure 7: Persistence in Time In-Between Robberies

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.

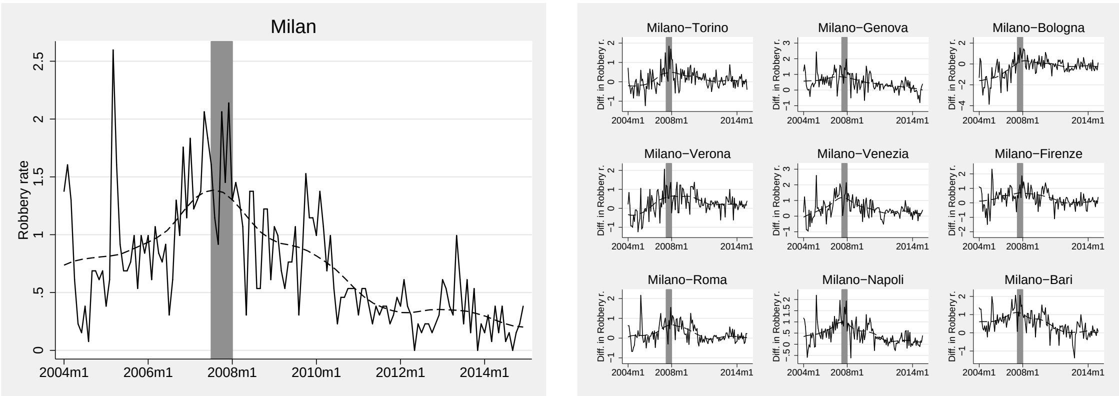

Figure 8: Bank Robbery Rates in Milan and in Comparison With Other Major Cities

Notes: In left panel the solid line represents the Milan monthly bank robbery rates per 100,000 inhabitants. The dashed line smoothes the solid line using a local linear regression. The right panel shows the differences in robbery rates (the raw data and a smoothed version) between Milan and the nine largest Italian cities.

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 (0/1)	0.14	0.35	0	1	0.13	0.34	0	1
Number of the sequence	5.10	6.88	1	84	4.20	5.53	1	84
Police (Polizia) Intervention (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 (0/1)	0.58	0.49	0	1	0.54	0.50	0	1
North-Western area (0/1)	0.35	0.48	0	1	0.38	0.48	0	1
North-Eastern area (0/1)	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 (0/1)	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 (×1000)	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 no firearm (0/1)	0.09	0.29	0	1	0.09	0.28	0	1
Some Italian involved (0/1)	0.79	0.41	0	1	0.77	0.42	0	1
Different nationalities (0/1)	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: 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*
	{*0.26*}	{*0.74*}	{*0.19*}	{*0.81*}
2009	0.139	0.128	0.060	0.180
	(0.348)	(0.335)	(0.239)	(0.385)
	72	*164*	*100*	*256*
	{*0.31*}	{*0.69*}	{*0.28*}	{*0.72*}
2008-2009	0.130	0.148	0.056	0.150
	(0.338)	(0.355)	(0.230)	(0.358)
	161	*420*	*161*	*513*
	{*0.28*}	{*0.72*}	{*0.24*}	{*0.76*}
2010	0.136	0.152	0.116	0.135
	(0.346)	(0.360)	(0.322)	(0.343)
	66	*158*	*129*	*288*
	{*0.29*}	{*0.71*}	{*0.31*}	*0.69*
2011	0.227	0.221	0.122	0.130
	(0.429)	(0.417)	(0.331)	(0.337)
	22	*77*	*41*	*131*
	{*0.22*}	{*0.78*}	{*0.24*}	{*0.76*}
2008-2011	0.141	0.157	0.088	0.143
	(0.348)	(0.364)	(0.283)	(0.350)
	249	*655*	*331*	*932*
	{*0.28*}	{*0.72*}	{*0.26*}	{*0.74*}

Notes: Standard deviations are shown in parentheses, the number of observations are shown in italics (fractions by police force are shown in curly brackets).

Table 3: Clearances by Number of the Sequence and Police Force

Number of the Sequence	*Carabinieri*	*Polizia*	*Polizia-Carabinieri*
1	0.130	0.148	0.017
	(0.338)	(0.355)	(0.032)
	161	*420*	
2	0.029	0.114	0.085**
	(0.171)	(0.320)	(0.043)
	34	*105*	
3	0.087	0.224	0.137*
	(0.288)	(0.419)	(0.077)
	23	*76*	
≥4	0.058	0.145	0.087***
	(0.234)	(0.352)	(0.031)
	104	*332*	

Notes: Years 2008 and 2009. Standard deviations (first two columns) and standard errors clustered by criminal group (last column) are shown in parentheses. The number of observations are shown in square brackets.

Table 4: Difference in Differences by Number of the Sequences and Police Force

	(1)	(2)	(3)	(4)	(5)	(6)
	The robbery has been cleared (0/1)					
Sample	*First robbery*		*Subsequent robberies*		*All robberies*	
Polizia Intervention	0.016	0.018	0.100***	0.087***	0.018	0.021
	(0.032)	(0.030)	(0.025)	(0.027)	(0.032)	(0.026)
Subsequent robberies					-0.078**	
					(0.032)	
Number of the sequence						-0.005***
						(0.002)
***Polizia* Intervention interacted with:**						
Subsequent robberies					**0.078***	
					(0.040)	
Number of the sequence						**0.009*****
						(0.003)
Constant	0.120***	-	0.074***	-	-	-
	(0.032)	-	(0.022)	-	-	-
Other Xs		√		√		
Observations	581	581	674	674	1,255	1,255
R-squared	0.001	0.209	0.020	0.101	0.009	0.010

Notes: Linear probability models with clustered (by criminal group) standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. 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 sequence.

Table 5: Difference in Differences by Same Day Robberies and Police Force

	(1)	(2)	(3)	(4)	(5)	(6)
	The robbery has been cleared (0/1)					
Sample	*Polizia in subsequent*		*Carabinieri in subsequent*		*Subsequent robberies*	
Polizia Intervention					-0.022	0.065*
					(0.057)	(0.035)
Different day	0.122***	0.129***	0.001	0.047	-0.011	0.026
	(0.031)	(0.038)	(0.061)	(0.075)	(0.056)	(0.051)
Number of the sequence						-0.001
						(0.003)
Number of the series × Different day robbery						-0.000
						(0.003)
***Polizia* Intervention interacted with:**						
Different day robbery					**0.133****	
					(0.064)	
Number of the sequence						**-0.006****
						(0.003)
Number of the sequence × Different day robbery						**0.012*****
						(0.004)
Constant	0.070**	-	0.059	-	-	-
	(0.031)	-	(0.066)	-	-	-
Other Xs		√		√	√	√
Observations	510	510	160	160	670	670
R-squared	0.020	0.120	0.000	0.290	0.030	0.034

Notes: Linear probability models with clustered (by criminal group) standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. The additional controls are listed in the notes of Table 4.

Table 6: Tighter Same Day vs. Different Day Differences

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	The robbery has been cleared (0/1)							
Robberies happen within	*any days*	*4 days*	*3 days*	*2 days*	*any days*	*4 days*	*3 days*	*2 days*
	Polizia				*Carabinieri*			
Different day	0.106***	0.109***	0.112***	0.107**	0.030	0.017	0.028	-0.037
	(0.026)	(0.032)	(0.035)	(0.045)	(0.044)	(0.048)	(0.053)	(0.047)
Constant	0.045*	0.045*	0.045*	0.045*	0.061	0.061	0.061	0.061
	(0.023)	(0.023)	(0.023)	(0.023)	(0.041)	(0.041)	(0.041)	(0.041)
Observations	929	406	317	213	330	136	101	76
R-squared	0.008	0.018	0.023	0.029	0.001	0.001	0.002	0.009

Notes: Linear probability models with clustered (by criminal group). All regressions control for four shift fixed effects (6 hour intervals). standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Regression Discontinuity Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	The robbery has been cleared (0/1)							
	Polizia				*Carabinieri*			
Post time of data update	0.090*	0.108**	0.107**	0.116***	0.029	0.047	-0.019	-0.021
	(0.047)	(0.046)	(0.046)	(0.044)	(0.057)	(0.061)	(0.050)	(0.053)
Time since data update (in hours)	-0.000	-0.001			-0.000	-0.001		
	(0.001)	(0.001)			(0.000)	(0.000)		
Constant	0.064**	0.060**	0.062**	0.060**	0.060	0.059	0.060	0.065
	(0.027)	(0.027)	(0.027)	(0.026)	(0.042)	(0.044)	(0.041)	(0.043)
Donut RDD (3 hours radius)		√		√		√		√
Time since data update is below	120 hours		32 hours		120 hours		32 hours	
Observations	489	466	248	225	163	157	82	76
R-squared	0.029	0.036	0.043	0.055	0.007	0.014	0.021	0.017

Notes: Linear probability models with clustered (by criminal group) standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Robustness Checks of Difference in Differences Estimates

	(1)	(2)	(3)	(4)	(5)
	The robbery has been cleared (0/1)				
Sample		*With CCTV coverage*	*Loot above average*	*No pharmacies*	*First daily robbery*
Model	*Probit*	*DinD*	*DinD*	*DinD*	*DinD*
Police Intervention	0.016	-0.010	0.002	0.010	0.003
	(0.030)	(0.050)	(0.035)	(0.040)	(0.039)
Subsequent robberies	-0.101**	-0.074	-0.032	-0.111***	-0.084**
	(0.045)	(0.046)	(0.037)	(0.041)	(0.037)
Polizia Intervention ×	0.106**	0.114*	0.081*	0.110**	0.095*
Subsequent robberies	(0.053)	(0.058)	(0.045)	(0.050)	(0.049)
Observations	1,255	702	642	787	746
R-squared		0.013	0.011	0.010	0.011

Notes: Linear probability model estimates (LPM) and probit marginal effects. Clustered (by criminal group) standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. All regressions control for year effects.

Table 9: Balance Test

	Polizia		Carabinieri		Polizia-Carabinieri	
	Average	SE	Average	SE	Average	SE
Cleared robbery	0.149	0.012	0.093	0.017	0.056	0.020***
Number of the sequence	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	0.160	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	0.198	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 and 2009. Standard errors are clustered by criminal group. For the last two columns only: *** p<0.01, ** p<0.05, * p<0.1.

Table 10: Selection and Simple Difference Estimates

	(1)	(2)	(3)
	The robbery has been cleared (0/1)		
Polizia Intervention	0.087***	0.090***	0.095**
	(0.028)	(0.025)	(0.038)
Polizia Intervention in $n-1$	0.028	0.035	0.041
	(0.033)	(0.028)	(0.036)
Polizia Intervention in $n-2$	-0.017		
	(0.032)		
Polizia Intervention in n and $n-1$			-0.007
			(0.051)
Observations	529	666	666
R-squared	0.021	0.023	0.023

Notes: Clustered (by criminal group) standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1. All regressions control for year effects.

Table 11: 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 12: 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 sequence	0.003**	0.004	0.009**	0.017***	-0.776***
	(0.002)	(0.004)	(0.004)	(0.005)	(0.198)
Loot > Loot average for the 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. 4. Standard errors are clustered by criminal group: *** p<0.01, ** p<0.05, * p<0.1.

Table 13: Difference in Differences in the Number of Bank Robberies

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Post-threshold used:	Jan-08					Jun-07	
Dependent variable:	Bank robbery rate (BRR)			Monthly change in BRR (CBRR)		BRR	CBRR
Post × Milan	-0.280***	-0.280***	-0.209**	-0.037***	-0.037***	-0.162***	-0.026***
	(0.028)	(0.028)	(0.089)	(0.003)	(0.003)	(0.029)	(0.003)
Milan	0.607***			0.019***		0.539***	0.013***
	(0.034)			(0.002)		(0.036)	(0.003)
Constant	0.307***	0.179***	0.617	-0.003***	-0.000	0.307***	-0.003***
	(0.019)	(0.000)	(0.472)	(0.001)	(0.000)	(0.019)	(0.001)
Month fixed effects	√	√	√	√	√	√	√
City fixed effects		√	√		√		
Region-specific trends			√				
Observations	12,804	12,804	12804	12,707	12,707	12,804	12,707
R-squared	0.051	0.132	0.145	0.011	0.011	0.050	0.011
Mean dep. var.	0.312	0.312	0.312	-0.00301	-0.00301	0.312	-0.00301

Notes: For each of the 97 major municipalities (*comuni capoluogo di provincia*) there are monthly observations on the number of bank robberies for a total of 11 years. In the first five columns the time of introduction of Keycrime is set to January 2008 (when the first predictions were used), in the last 2 columns it is to June 2007 (when the data gathering process started). Clustered standard errors at the municipality level in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

A Online Appendix

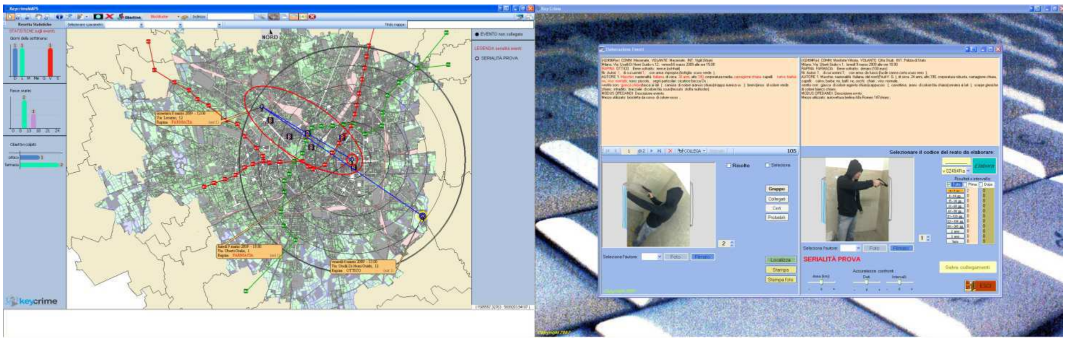

Figure 9: Comparison of Events

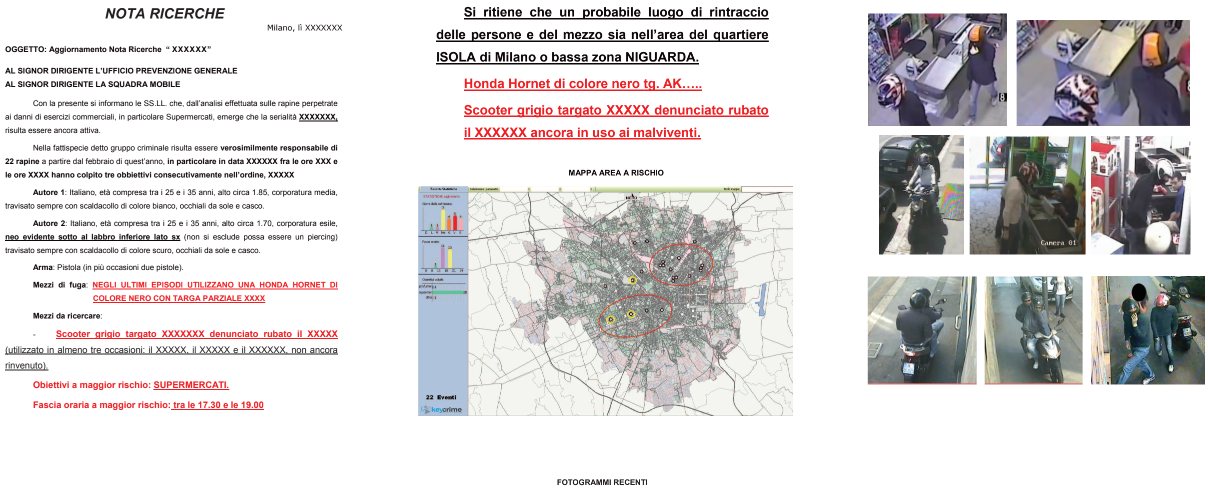

Figure 10: Instructions for Police Patrols

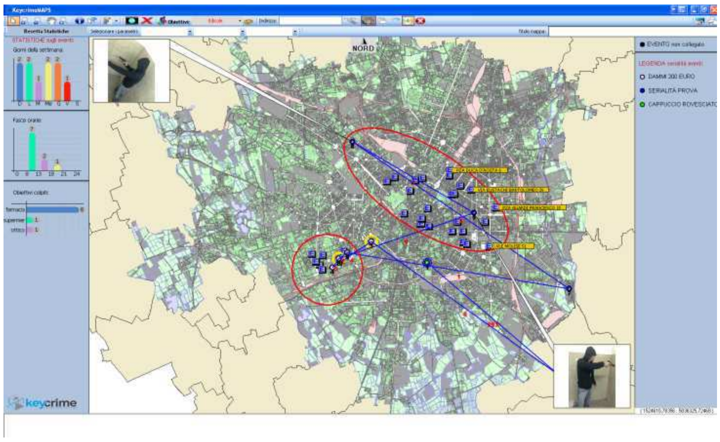

Figure 11: Predicted Targets

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

Figure 12: Gendarmerie and Police

Table 14: Balance Test

	1st Strategy		2nd Strategy	
	Average	SE	Average	SE
Cleared robbery	0.077	0.040*	0.112	0.035***
North-Western area	0.010	0.067	0.004	0.069
North-Eastern area	-0.002	0.049	-0.060	0.064
Year	-0.065	0.067	0.057	0.163
Month	0.736	0.500	-0.012	0.609
Day of the month	0.921	1.113	1.474	1.204
Sunday	0.009	0.033	0.009	0.040
Monday	0.041	0.053	-0.166	0.066**
Tuesday	0.011	0.046	0.029	0.038
Wednesday	-0.002	0.044	0.027	0.042
Thursday	-0.047	0.043	0.028	0.047
Friday	-0.054	0.054	0.061	0.039
Saturday	0.042	0.048	0.013	0.046
Daylight	0.057	0.064	0.067	0.073
Average age	-2.115	1.927	-0.280	1.412
Amount stolen in euros	0.620	1.343	0.070	0.747
Firearm 0/1	0.021	0.053	0.019	0.044
At least one knife, but no firearm	0.007	0.034	-0.002	0.064
Some Italian	0.027	0.059	-0.060	0.037
Different nationalities	0.034	0.040	0.046	0.040
Number of robbers	0.044	0.089	0.063	0.101
Pharmacy	0.008	0.067	0.204	0.070***
Other business	-0.053	0.042	-0.098	0.048**
Supermarket	-0.011	0.056	-0.081	0.082
Bank	0.064	0.041	-0.011	0.032
Video rental	-0.038	0.028	0.027	0.022
Tobacco shop	0.012	0.014	-0.018	0.026

Notes: The 1st Strategy coefficient is the difference-in-differences between *Polizia* and *Carabinieri* for subsequent and first event. The 2nd Strategy coefficient is the difference between the robberies investigated by the *Polizia* that happen one day later and those that happen on the same day. Standard errors are clustered by criminal group. For the last two columns only: *** p<0.01, ** p<0.05, * p<0.1.

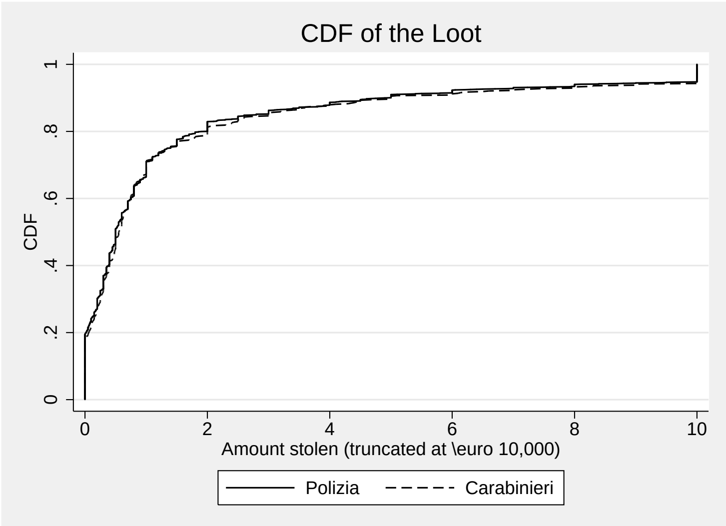

Figure 13: Cumulative Distribution of the Loot

Notes: The loot is expressed in €1,000 and is truncated at 10,000 to focus where most the data are. The Kolmogorov-Smirnov test for equality of distribution functions cannot reject that null that the distributions are the same (irrespective of truncation).