PAVED with good intentions ? An evaluation of a French police predictive policing system

Yann Lecorps* Gaspard Tissandier†

Work in progress

For the latest version of the paper, please click here

Abstract

From late 2017 to early 2019, one of the two french law enforcement agencies tested in 11 out of 101 departments a predictive policing system named PAVED. The system design by the *Gendarmerie* predicts burglaries and vehicle thefts with the stated objective of better allocating patrols and thus increasing deterrence. We use month-law enforcement jurisdiction area panel data to evaluate whether the system produces the expected reduction in these thefts and whether this effect is due to a deterrent effect or a displacement effect. Both the standard Two Way Fixed Effect and Synthetic Difference-in-Difference estimations consistently indicate a significant reduction of vehicle thefts in the treated *Gendarmerie* areas but no detectable effect on burglaries. Most of our results indicate that vehicle theft has not increased in areas near the treated areas, suggesting that the reduction in vehicle theft is due to a deterrent effect rather than a displacement effect.

Keywords : predictive policing, paved, deterrence, law enforcement
JEL code : K14, K42, O17

We thank Philipp Ketz for his valuable comments during the Doctorissmes conference at Paris School of Economics (PSE), Dan O'Flaherty for his amazing advising, and the participants of the Public Policy Seminar (S2P) at the University of Paris 1 for their helpful discussions.

*Université du Mans (GAINS), yann.lecorps@univ-lemans.fr

†Centre d'Économie de la Sorbonne (CES), Université Paris 1 Panthéon Sorbonne, gaspard.tissandier@univ-paris1.fr

1 Introduction

Since Machine Learning methods started to pertain from standard applications such as computer vision or pattern recognition, the usages of prediction systems in public policy and social science have increased rapidly. One of the most debated applications of Machine Learning in public policy settings is predictive policing. The objective of predictive policy systems is to anticipate crimes in order to guide police patrol and better allocate law enforcement units. In France, the law enforcement service dedicated to rural and small urban areas, the *Gendarmerie*, has developed in 2016 a predictive policing system called PAVED[1]. In one of the few public expressions of the *Gendarmerie* about PAVED, Colonel Collorig clearly states the deterrent objective of this new tool:[2]: *"Our job is not to let an offense be committed but to prevent it from being committed. It's not software that's designed to catch offenders, it's software that's designed to secure the population and our jurisdiction."*. While the effects on crime of some other predictive policing applications such as "KeyCrime" in the city of Milan, or PredPol in a US agglomeration, have been assessed and are promising (Mastrobuoni, 2020 and Jabri, 2021), the effects of the predictive system PAVED of the French *Gendarmerie* remain an open empirical question.

Although little information is made public by the *Gendarmerie* about this algorithm, we know that PAVED has been tested in 11 out of 101 French departments over a period of several months between 2017 and 2019 to predict burglaries and vehicle thefts[3]. The application was consultative by any *gendarme* from a device connected to the intranet of the *Gendarmerie* with the objective to make the prediction accessible for field units but without substituting to *gendarmes*'s analytic skills. There may be heterogeneity in the way PAVED was used by *gendarmes* in the 11 test departments although there is a long and widespread culture of heatmap use by this agency (Germes, 2014). This test phase was to be followed by a national deployment of the application and the prediction was scheduled to

[1]PAVED stands for: *" Plateforme d'Analyse et de Visualisation Évolutive de la Délinquance"*
[2]Full podcast available here: https://www.radiofrance.fr/franceculture/podcasts/la-methode-scientifique/22-v-la-la-police-predictive-1641835
[3]In September 2019 the newspaper Mediapart published a memo from the Central Criminal Intelligence Service (SCRC) of the *Gendarmerie* available at: https://www.mediapart.fr/journal/france/020919/la-gendarmerie-ne-parvient-toujours-pas-predire-l-avenir 2

expand to other crimes. However, according to an interview with Colonel Perrot conducted by Nano and Tréguer (2022), PAVED has apparently not been generalized beyond the test phase.

We use this quasi-experimental setting to estimate the direct effect of PAVED on burglaries and vehicle thefts rates as well as potential geographical displacement or diffusion effects. We base our analysis on monthly records of thefts at the level of the *Gendarmerie* area, *Police* area, and entire departments over the period 2015-2018. The map in Figure 9 depicting the average monthly rate of burglary and vehicle theft by departments over the period 2015-2018 provides an initial visualization of the geographical distribution of our dataset.

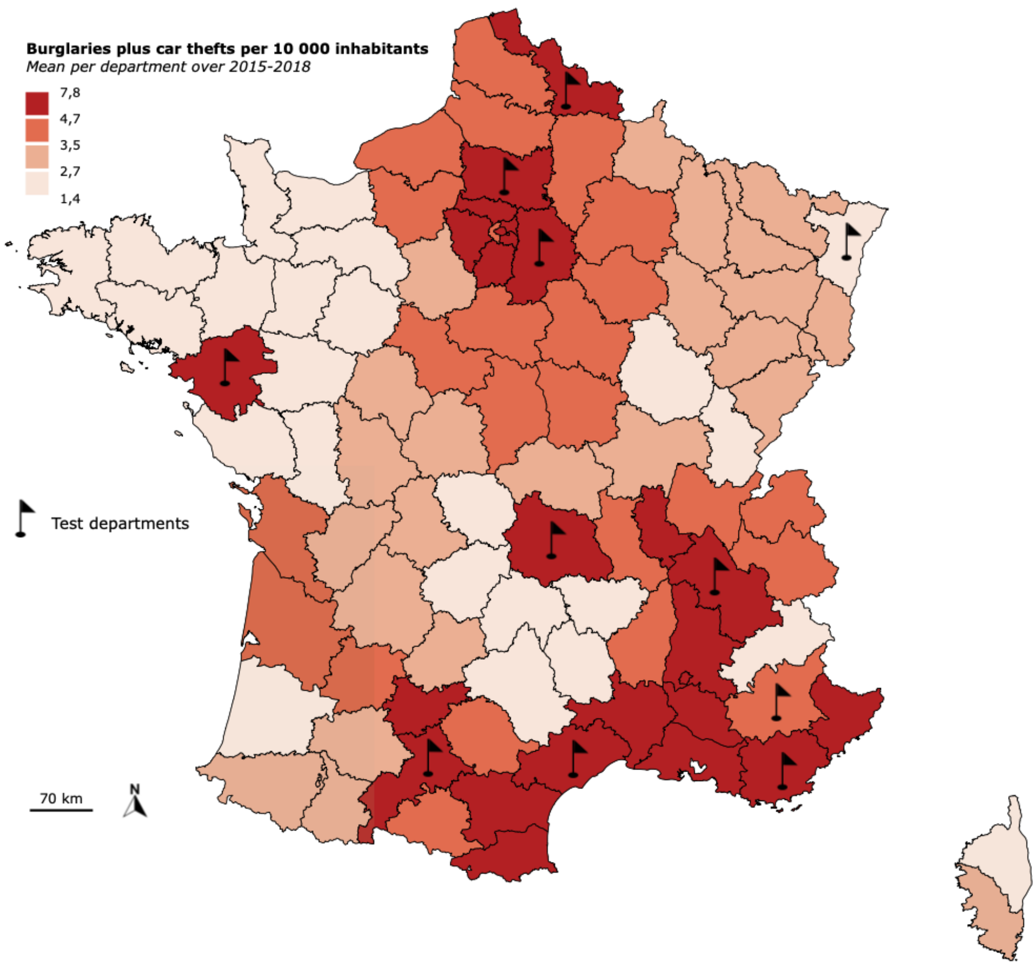
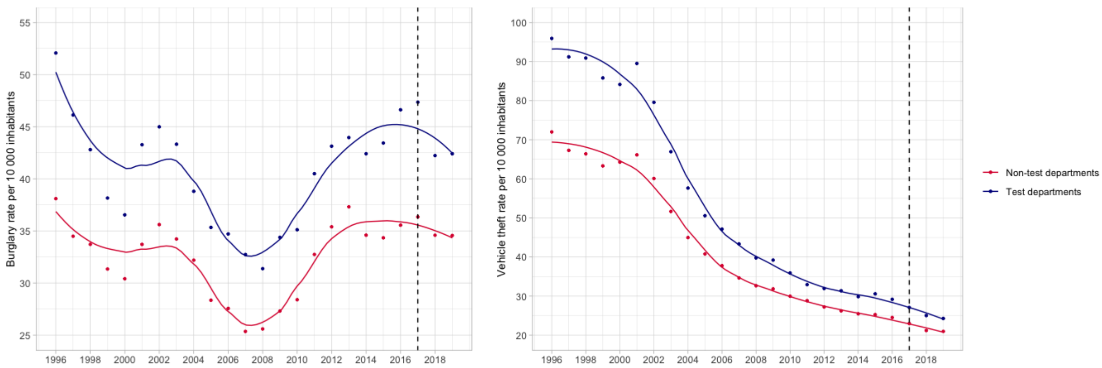


Figure 1: Burglaries plus vehicle theft rate per 10 000 inhabitants

We observe disparities ranging from 1 to almost 6 in the rate of burglaries and vehicle thefts between French departments. The southeast of France (the French Riviera) and the Paris region are the two most affected areas, whereas the northwest of France is much less affected. We also note that 9 of the 11 test departments are in the fourth quartile (between 4.7 and 7.8 burglaries and vehicle thefts per 10,000 inhabitants on average per month) suggesting not surprisingly that the level of burglary and vehicle theft rates has been an important criterion used by the *Gendarmerie* to select test departments.

Figure 2: Evolution of burglary and vehicle theft rates in France, 1996-2019



Note: Each point measures the burglary and vehicle theft rate per year for the groups of test and non-test departments. The curves are obtained by local polynomial smoothing (loess).

The graphs in Figure 2 provide an overview of the evolution since 1996 in the rates of burglary and vehicle theft in departments where PAVED has been tested between 2017-2019 and those where it has not. These two graphs are consistent with the map in Figure 1 as burglary and vehicle theft rates are consistently higher in tested departments than in non-tested departments over two decades. The very different patterns observed between burglary and vehicle theft suggest that they were not similarly affected by the events or their environment over the period. There has been a significant and continuous decline in the rate of vehicle theft since 1996, from 95 to 25 for tested departments and from 70 to 20 for non-tested departments, with a narrowing gap between tested and non-tested departments. Burglaries reached their lowest point in the late 2000s before increasing for just over a decade and then declining again to a rate of about 40 per 10,000 inhabitants in the tested departments and 35 in the non-tested departments. Given these significant differences in pattern, we deal separately with the effect of PAVED on burglary and vehicle theft in the remainder of the paper.

As pointed out by Durlauf, Navarro, and Rivers (2010), aggregate regressions are demanding in statistical hypotheses in order to generate a causal relationship between the treatment and the outcome. To identify the treatment effect of PAVED, we rely on two Differences-in-Differences methodologies. The first is an adapted version of the Two-Way Fixed Effects (TWFE) model which accounts for the seasonal structure of departmental criminality. The second method is the Synthetic Difference-in-Difference (SDID) as proposed by Arkhangelsky et al., 2021. This method generates a synthetic counterfactual respecting the parallel trend assumptions in the pre-treatment period by creating a weighted control group in geographical units and time units in all the observations.

The objective of the paper is to identify the direct and indirect effects of PAVED on targeted criminality (burglary and vehicle theft). The first estimation uses a TWFE model on two separate areas: the *Gendarmerie* jurisdiction over the whole country, and the *Police* jurisdiction over the whole country. The intuition behind this split is to include similar jurisdiction in the control and treatment groups. With the TWFE model, the choice of the control group is straightforward as we include all the departments which are not either a treated department or a nearby department, but it could be the case that some departments have more of their place than others in this control group. By using the SDID, we test if choosing precisely the control group leads to a different estimation of the effect of PAVED on the treated departments. Finally, we test for the direct and indirect effect of PAVED by considering at the same time all French areas, both *Gendarmerie* or *Police*, by estimating an augmented TWFE model. This final step helps to disentangle the potential displacement effect of criminality due to PAVED, across all potential neighboring areas.

We find consistent results across the different specifications employed. Our results indicate that PAVED significantly reduces vehicle theft in the *Gendarmerie* areas where it was tested : vehicle theft per 10,000 inhabitants decreased by 5 to 3 percent depending on the specification. This effect corresponds to an average reduction per department and year in the treatment group of around 114 to 68 vehicle thefts. We also find a reduction effect at the department level although less important in magnitude and significance. Conversely, burglary rates appear unaffected by the use of this system. The absence of a significant effect of PAVED on burglary rate is found at all levels of law enforcement jurisdiction : *Gendarmerie*, *Police*, and the entire department[4]. Finally our most robust estimates does not indicate an increase in the rate of vehicle theft in areas near the treated areas, suggesting that the reduction in the treated areas is due to a deterrent effect rather than a displacement effect.

The remainder of the paper is organized as follows: section 2 presents the challenges of estimating deterrence and displacement in the institutional context surrounding the implementation of PAVED, section 3 presents the results of the different specifications and section 4 concludes.

2 Estimating the potential deterrent and displacement effects of PAVED on crime

2.1 Framing the research question

We want to know whether or not PAVED affects the level of two types of theft (burglary and vehicle theft) over the test period. Formulated into an equation, this can be written as:

$$\text{Theft}_{i,t} = \alpha + \beta \text{PAVED}_{i,t} + \epsilon_{i,t} \tag{1}$$

where $Theft_{i,t}$ is the level of burglary or vehicle theft in the area of the department i in month t and β the average effect of PAVED in departments where the algorithm has a potential effect in tested months, and $\epsilon_{i,t}$ the residual term. We believe that β, the effect of PAVED, can be of two different natures. During the implementation of PAVED, we can expect a reduction in the level of theft in the treated areas through deterrence while in the areas close to the treated areas we can expect an increase in theft through displacement. To decompose these

[4] The sum of the *Gendarmerie* and *Police* areas covers the whole department.

two potential effects, we create the binary variables $Post_t$ equal to 1 during the treatment period and 0 otherwise, $Treatment_i$ equal to 1 for treated areas, and 0 otherwise and $Catchment_i$ equal to 1 for areas near the treated areas and 0 otherwise. To be considered in the catchment area, a zone has to be close geographically to a treated area. By interacting these variables, the potential deterrent effect of PAVED is captured by the combination ($Treatment_i \times Post_t = 1$), and the potential displacement effect is captured by the combination ($Catchment_i \times Post_t = 1$). Equation thus 1 rewrites:

$$\text{Theft}_{i,t} = \alpha + \beta_1(\text{Treatment}_i \times \text{Post}_t) + \beta_2(\text{Catchment}_i \times \text{Post}_t) + \epsilon_{i,t} \qquad (2)$$

With β_1 the average marginal deterrent effect of PAVED in the treated areas and β_2 the average marginal displacement effect of PAVED in catchment areas. We expect β_1 to be negative and β_2 to be positive. If PAVED produces a pure marginal deterrence effect in the treated areas without displacement in the neighboring areas, β_1 may be negative and β_2 not significantly different from 0. However, if β_1 is not significantly different from 0, we expect β_2 to be not different from 0 either. The unobserved variables that affect crime differently in each area each month are captured by $\epsilon_{i,t}$. Below we discuss the institutional features surrounding the implementation of PAVED that may affect the true magnitude of β_1 and β_2 as well as the methodological challenges that may affect these magnitudes for the wrong reasons.

2.2 Magnitude of β_1: deterrent effect

As indicated in the introduction, little information is publicly available and despite multiple requests on our part we have not obtained any further information from the *Gendarmerie* on PAVED for this analysis. However, thanks to various media supports and a report of Nano and Tréguer (2022)[5], we know several essential features of PAVED implementation. It was created by a national service of the *Gendarmerie* called *"service central de renseignement criminel de la gendarmerie*

[5]We thank Félix and Edlira for the mutual exchanges about PAVED and for giving us access to their report.

nationale" to orient units toward areas more likely to host crime to maximize deterrence. The software uses a Machine Learning model to predict burglary and vehicle theft evolution in France and to generate "heatmaps" on a local level (neighborhood level), many times per day[6]. Specifically, it uses *police* and *Gendarmerie* records of crimes and various socio-demographic indicators to predict where and when future burglary and vehicle theft are likely to occur.

In a Beckerian approach (Becker, 1968, Polinsky and Shavell, 2000), a predictive policing application is indeed expected to generate a positive productivity shock for law enforcement. Predicting the areas and times crime is most likely to occur should allow law enforcement to be in the "right place at the right time" to deter. More precisely, this potential additional deterrence is generated by the increased accuracy of law enforcement presence in the field, which increases the probability of detection of criminals. As a result of this increase in the probability of detection, the crime is no longer sufficiently rewarding for some of the individuals and they renounce committing the crime ($\downarrow \beta_1$). There is a fairly extensive empirical literature consistent with this theoretical expectation that, given constant resources, in various contexts redeploying law enforcement to hot spots can reduce crimes through deterrence (see Braga, 2008 and Chalfin and McCrary, 2017).

In the specific context of the implementation of PAVED in France, several points are likely to affect the magnitude of β_1. First, the analysis focuses on two different types of theft that can be affected differently by PAVED. Regarding similarities, burglary and vehicle theft are both property crimes with primarily pecuniary motivations (Apel and Nagin, 2011) and they are analyzed together in various literature on crime deterrence (see for example Levitt, 1997 and Lochner, 2004). Under French law, they are similarly considered as "theft with the aggravating circumstance of breaking and entering". They are punished with an identical maximum sentence of 7 years in prison and a 100,000 euro fine[7]. This is an important point because the severity of punishment is a key parameter of deterrence and

[6] According to the sources, PAVED provides prediction for two periods within each day

[7] The sentence can be more severe in case of additional aggravating circumstances during the commission of the burglary or vehicle theft (violence, violence leading to death, etc). French criminal law regarding theft available at: `://legifrance.gouv.fr/codes/section_lc/LEGITEXT000006070719/LEGISCTA000006165324`

can affect the magnitude of β_1. This common upper bound on sentencing makes it implausible that a difference in severity could explain a difference in the marginal effect of PAVED on burglary and vehicle theft.

Despite the proximity in many respects of these two crimes, it is arguable that PAVED does not necessarily produce a symmetric deterrent effect. One of the critical links in the perceptual deterrence chain (Apel, 2013) is police visibility. Since *gendarmes*' patrols occur on the streets and vehicle thefts also occur on the street, one might think that mutual visibility between *gendarmes* and criminals is higher for vehicle thefts than for burglaries. This difference in visibility could lead to a difference in the perception of risk among criminals, resulting in a difference in magnitude of β_1. Another point is that β_1 captures a marginal effect whose magnitude depends on the *gendarmes*' initial level of knowledge of their territory. The lower the level of knowledge of the *gendarmes* about their jurisdiction, the higher the marginal deterrent effect of PAVED can be expected to be. By definition burglaries concern immobile goods, whereas vehicle thefts concern mobile goods. Detailed knowledge of vehicle thefts in a territory is therefore perhaps more difficult to obtain on vehicle thefts than burglaries. If this is the case, one can expect a higher added value and thus a higher marginal effect of PAVED on vehicle thefts than on burglaries.

Finally, PAVED potentially modifies the opportunity cost between treated and untreated areas which can lead to a displacement of crime. Recall that our goal is to capture a β_1 which is a deterrent effect, not a reduction in theft levels which could have an alternative explanation such as displacement in areas close to the treated areas. If a displacement effect occurs following the introduction of PAVED, the associated criminality will decrease in the treated departments and increase in the neighboring department. Here, we are exposed to the risk of capturing a β_1 whose magnitude is overestimated for reasons unrelated to deterrence if we include those neighboring departments in the control group. We, therefore, introduce the interaction variable $(Catchment_i \times Post_t)$ in the equation 2 with β_2 capturing the potential displacement effect in areas close to the treated areas. The institutional context surrounding PAVED that may affect the magnitude of β_2 is discussed in the next section.

2.3 Magnitude of β_2 : displacement effect

As noticed previously, the implementation of PAVED in the *Gendarmerie* areas of the tested departments is likely to create a displacement effect of burglary and vehicle theft in nearby untreated *Police* and *Gendarmerie* areas. The literature on this topic is mixed. Depending on the context and the type of crime, it seems that the redeployment of resources may generate an almost complete displacement effect (Ho, Donohue, and Leahy, 2013) or no displacement effect at all and even a positive spillover effect (Weisburd et al. (2006), Guerette and Bowers (2009)). On this point, the organization of law enforcement in France is crucial.

Like numerous other countries, France has two distinct law enforcement agencies: the *Police* and the *Gendarmerie*. *Police* is dedicated to cities and urban areas, while the *Gendarmerie* is dedicated to rural and small urban areas[8]. Except for a few services, the mission of the *Police* and the *Gendarmerie* are similar, and organized similarly. For both, the main missions are criminal police and administrative police. The criminal police missions are to ascertain crimes and misdemeanors, gather evidence, and find perpetrators. The administrative police missions are mostly to prevent public order offenses. From a global point of view, the mission of criminal police is repression while the administrative police are intended to deter criminality. Although other marginal differences characterize these two forces, they are primarily distinguished by separate geographic areas of jurisdiction. The *Gendarmerie* jurisdiction area represents 95% of the French territory and 50% of the population. PAVED has been developed by the *Gendarmerie* at a national level but tested at the department level.

These jurisdictional distinctions provide two valuable insights regarding the potential magnitude of β_2. First, the use of PAVED in treated departments has no reason to affect the number and deployment strategies of *gendarmes* in untreated departments. Second, the use of PAVED in the *Gendarmerie* zones of the treated departments has no reason to affect the number and deployment strategies of the *Police* in the *Police* urban areas. This implies that the level of deterrence in control

[8]To be more precise, the population threshold in inhabitants per city between *Police* and *Gendarmerie* is around 20 000 inhabitants: for cities with 20 000 or more inhabitants, the *Police* is qualified and the *Gendarmerie* is qualified for all other areas. See : *Loi n°95-73 du 21 janvier 1995 d'orientation et de programmation relative à la sécurité.*

areas is independent of the implementation of PAVED in treated areas. This makes it credible that the estimates of β_1 and β_2 are entirely attributable to PAVED and not partly affected by an underlying and uncaptured change in the allocation of law enforcement resources between treated, catchment, and control areas.

2.4 Data

We use an open access database from the French Ministry of the Interior providing the number of burglaries and vehicle thefts per department and per month since 1996[9]. This month-department panel data is made up of the data collected by the *Gendarmerie* in the *Gendarmerie* areas and the data collected by the *Police* in the *Police* areas. The level of crime in the department is given by the sum of the *Gendarmerie* and *Police* areas. This database provides a reliable measure of burglary and vehicle theft levels since the reporting rates for burglary are over 70% and over 90% for vehicle theft[10]. Specifically, we construct our measure of burglary as the sum of the categories of burglary of primary and secondary homes while our measure of vehicle theft is the sum of theft of cars, motorcycles, and commercial vehicles.

Several constraints lead us to reduce the size of the panel used to conduct the empirical analysis. First, in 2015, a break in the way burglaries are recorded does not allow us to ensure the reliability of the comparison between pre-2015 and post-2015 data[11]. On the temporal aspect of the panel, we do not know with certainty in which month of 2019 the experimentation stopped in the 11 test departments. Therefore we limit the panel to the month of December 2018. Finally, Paris and three departments[12] on the periphery of Paris do not have a *Gendarmerie* area. We remove them from the sample for convenience in the econometric analysis in which we measure potential displacement effects between *gendarmerie* and *Police*

[9] Available at: https://www.data.gouv.fr/fr/datasets/chiffres-departementaux-mensuels-relatifs-aux-crimes-et-delits-enregistres-par-les-services-de-police-et-de-gendarmerie-depuis-janvier-1996/

[10] See: Enquêtes Cadre de vie et sécurité 2010 à 2019 : https://www.interieur.gouv.fr/Interstats/L-enquete-Cadre-de-vie-et-securite-CVS

[11] See at the link of the previous note the *"documentation des chiffres mensuels départementaux"* file for more details.

[12] Haut de Seine (92), Seine Saint Denis (93) and Val de Marnes (94)

areas within the same department. We are therefore working on a data panel of 92 departments observed monthly over the period 2015-2018. In order to limit the impact of the overall decreasing trend in criminality observed in the introduction, we limit the pre-treatment period from 08/2016 to 08/2017, while the treatment period spans from 09/2017 to 12/2018[13].

As described before, we define three different types of areas: treatment, catchment, and control. The treated areas are those where PAVED was tested, the catchment areas are defined as sharing a common geographical border with the treated areas. The control areas include all others areas. We have 92 departments in the database, with 11 departments in the treatment group, 38 departments in the catchment group, and 43 departments in the control group. Within each department we distinguish two areas: *Gendarmerie* and *Police*. This distinction will be used to precisely identify the effect of PAVED on treated areas, and on the potential catchment areas.

As the areas studied (*Police*, *Gendamerie*, and the entire department) have very heterogeneous populations, we match each observation at the area-month level with its population for the rural (*Gendarmerie*) area, the urban (*Police*) area and the total population in the department. The population levels are imputed by using the communal census database[14]. We work on outputs that are rates of burglary and vehicle theft per 10,000 inhabitants.

2.5 Empirical Strategy 1 : Two-Way-Fixed-Effects (TWFE)

In order to adequately capture the potential effects described above, we need to reduce the threats of endogeneity that would lead to biased estimates of β_1 and β_2. As discussed and as shown in the preliminary representations of our outputs in the introduction, we are primarily exposed to treatment group selection bias and omitted variable bias. To reduce this threat, a first step is to augment the specification of equation 2 with two-way fixed effects such as:

[13] The 11 test departments are: Alpes de Haute Provence (04), Haute-Garonne (31), Hérault (34), Isère(38), Loire Atlantique (44), Nord (59), Oise (60), Puy de Dôme (63), Bas-Rhin (67), Seine-et-Marne (77) and Var (83)

[14] Census of the French population produced by the French Institute of Public Statistics (INSEE) available at: `https://www.insee.fr/fr/statistiques/3698339`

$$Y_{i,t} = \alpha_i + \phi_t + \beta_1 \left(\text{treatment}_i \times \text{post}_t\right) + \beta_2 \left(\text{catchment}_i \times \text{post}_t\right) + \epsilon_{i,t} \tag{3}$$

With i the department index, t an indicator of the time period (month $\times$ year). Here, α_i captures the combined effects of the differences between departments that are invariant over time and ϕ_t captures the combined effect of variables which changes over time but not departments. The outcome $Y_{i,t}$ measures the rate of burglary or vehicle theft per 10,000 inhabitants per department and month, to which a $log + 1$ transformation is applied to facilitate interpretation of the coefficients. Equation 3 describes a standard Two-Way-Fixed-Effect specification, relying on the Difference in Differences (DiD) framework.

An important issue regarding TWFE model is at which scale should we introduce the fixed effect. In France, we observe many nested geographical levels : department are composed of many cities, and each department are included into a region. Each region groups from 2 to 12 departments, with an average of 7 departments per region. As seen in Figure 3 and 4, the criminality disparity by region is strong, as is the spread of criminality. We observe that the median rate of burgalry per region is higher than the vehicle theft rate, which is confirming the observation of Figure 2. Regarding the burglary rate, the highest median is the Provence-Alpe-Côte d'Azur (PACA), while the lowest is the Pays de la Loire (PL). However, those two regions exhibits outliers, with the lowest observations of PACA almost equal to the PL median. The spread of the interquartile range also indicate a strong heterogeneity of the department within each region, as is the presence of numerous outliers for 5 out of 13 regions.

The spread of the interquartile range is less important for the vehicle theft rate, but the presence of many outliers (in 4 regions) also indicate a strong heterogeneity of this rate across department. While some region are still highly criminalistic, as PACA or Ile de France (IDF), some region doesn't follow a clear pattern like Auvergne-Rhône-Alpe which exhibits a high median regarding bulgary rate, but in average median regarding the vehicle theft rate.

These observations motivate us to introduce departmental scale geographical fixed effect, as the region ones cannot guarantee an effective control of the hetero-

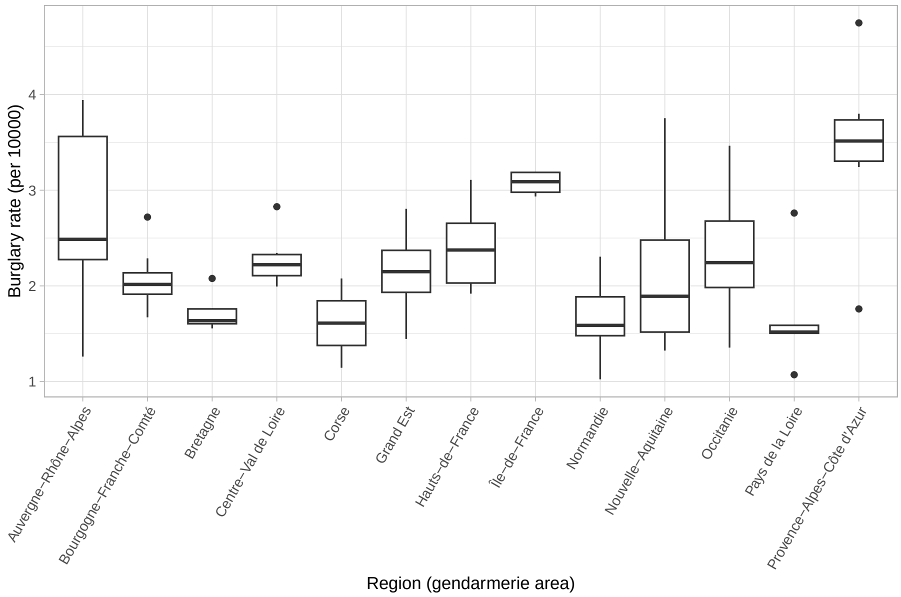


Figure 3: Burglary rates by region (2018)

Note: Each box represents the distribution of the burglary rate within region. The thick line in the box is the median, the extremities of the box are the first (lower) and the third (higher) quartiles. The difference between the bottom and the top of the box is defined as the interquartile range, and the end points of the line represents the maximum and minimum, computed as Q_1 - 1.5 IQR and Q_3 + 1.5 IQR. All the single points are outliers : they are outside of the minimum-maximum interval

geneity of department within each region. However, we present the results using consecutively regional and departmental fixed effects.

Looking at the plots in Figures 5 and 6, which represent the monthly rates of burglary and vehicle theft in a subsample of departments, one notices a fairly clear seasonality in the series that is specific to each department. In the department of *Rhône* (69), vehicle thefts are higher each year around October, whereas in the department of *Pyrenees Orientales* (38), located on the Mediterranean See, the trend is different, with vehicle thefts tending to be highest during the summer. In order to account for this heterogeneity for each department for different time periods, we augment equation 3 with a month-department fixed effect $\delta_{i,m}$ such that :

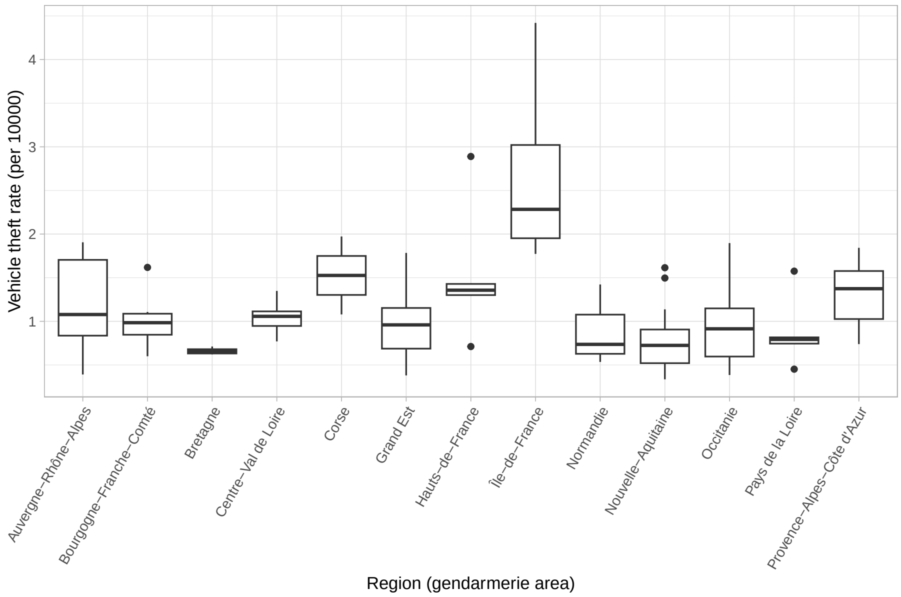


Figure 4: Vehicle theft rates by region (2018)

Note: Each box represents the distribution of the vehicle theft rate within region. The thick line in the box is the median, the extremities of the box are the first (lower) and the third (higher) quartiles. The difference between the bottom and the top of the box is defined as the interquartile range, and the end points of the line represents the maximum and minimum, computed as Q_1 - 1.5 IQR and Q_3 + 1.5 IQR. All the single points are outliers : they are outside of the minimum-maximum interval

$$\log\left(\frac{c_{kt}}{pop_{kt} \times \frac{1}{10{,}000}}+1\right) = \alpha_i+\phi_t+\delta_{i,m}+\beta_1\,(\text{treatment}_i \times \text{post}_t)+\beta_2\,(\text{catchment}_i \times \text{post}_t)+\epsilon_{i,t} \tag{4}$$

With k the type of crime, i the department index, t an indicator of the time period (month $\times$ year), m the month (January, February,...). The fixed effect $\delta_{i,m}$ accounts for the heterogeneity of the department with respect to each month : if a department i have structurally more criminality during summer, and a department j during winter, δ will account for it.

Since the main focus of this paper is the effect of PAVED on the *Gendarmerie* areas, we present plots of the criminality level of this jurisdiction. The plots for *Police* and whole department are to be found in figure 11, 12 and in appendix, section 4.1. Plotting the linear trends associated with the treatment and control group for burglary and vehicle theft in the graphs in Figures 7 and 8 provides a check on the suitability of applying the TWFE method to our setting. We again verify that the rates of burglary and vehicle theft are higher in the treatment group than in the control group. For burglary, the pre-treatment trends seem quite convincing and something different seems to happen post-treatment as the treatment group trend decreases slightly faster than the control group trend. For vehicle theft, the pre-treatment trends are less convincing, since a steeper linear downward trend is observed in the treatment group than in the control group. Post treatment, the downward trend continues with a slightly steeper slope for the control group.

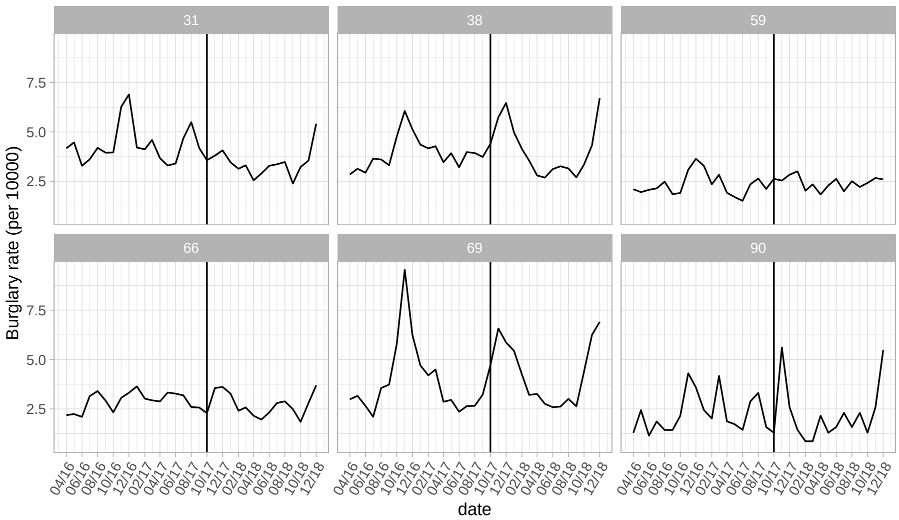


Figure 5: Burglary rates from January 2016 to January 2019 for 6 departments

Note: Each plots represents the burglary rate of 6 randomly drawn departments. In these plots, the treated departments are : Haute-Garonne (31), Isère (38), Nord (59). The control departments are : Pyrénées-Orientales (66), Rhône (69), Territoire de Belfort (90)

The graphs in Figures 7 and 8 do not provide clear preliminary insight regarding the potential effect of PAVED in the treatment areas. They provide a fairly convincing rationale for the applicability of TWFE to our research question although some doubts may remain. In order to dispel as much as possible the doubts that might remain, we rely on a second empirical strategy to improve the comparability between the treatment and control groups. We describe this strategy in the next section.

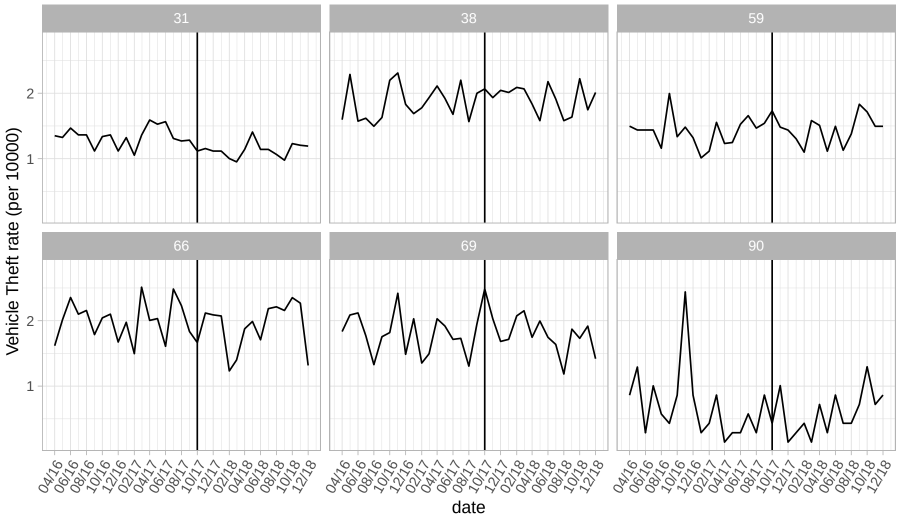


Figure 6: Vehicle theft rates from January 2016 to January 2019 for 6 departments

Note: Each plots represents the vehicle theft rate of 6 randomly drawn departments. In these plots, the treated departments are : Haute-Garonne (31), Isère (38), Nord (59). The control departments are : Pyrénées-Orientales (66), Rhône (69), Territoire de Belfort (90)

2.6 Empirical Strategy 2: Synthetic Differences in Differences

The key assumption of TWFE model is the outcomes parallel trends : both control and treated group outcomes would have to follow the same trend if the treatment wasn't applied. While this hypothesis is not testable per se, as the treatment was effectively administered, we can observe the pre-treatment trends of both groups to reinforce the confidence in the choice of the control group. As seen in the introduction, the treatment group departments are widely distributed in France, but not homogeneously distributed across the country. Thus, some department may be more suitable for being included in the control group than other, regarding various characteristics such as the geographic placement, or other factors. This is on this intuition that Arkhangelsky et al., 2021 build the Synthetic Difference-in-Difference method in order to adequately choose units and periods in the control group. The objective is to re-weight unit and time period in the control group in order for it to match as closely as possible the pre-trend from the treatment group, thus obtaining a suitable control group for assessing the effect of PAVED.

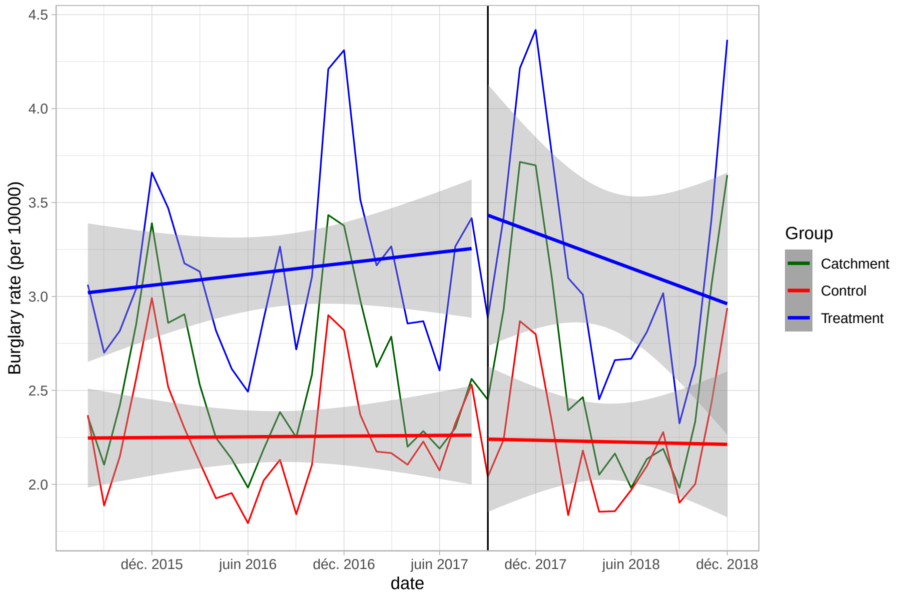


Figure 7: Burglary rates per month from January 2016 to January 2019

Note: this plot represents the trend of burglary rate for the treatment, catchment and control group. Four trend lines are fitted, during the pre-treatment and during the treatment. Two are fitted for the treatment group and two for the control group.

To align pre-exposure trends of the control and treatment group, we find weights such as : $\sum_{i=1}^{N_{co}} \hat{\omega}_i^{sdid} Y_{it} \approx N_{tr}^{-1} \sum_{i=N_{co+1}}^{N} Y_{it}$. Then, these weights are used in a Difference-in-Difference estimation :

$$(\hat{\tau}, \hat{\mu}, \hat{\alpha}, \hat{\beta}) = \text{argmin} \left\{ \sum_{i}^{N} \sum_{t}^{T} (Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2 \hat{\omega}_i^{sdid} \hat{\lambda}_i^{sdid} \right\} \tag{5}$$

The geographic unit weights ω_i are obtained by minimizing the following objective function. The time unit weights λ_t are found similarly (see Arkhangelsky et al., 2021).

$$(\hat{\omega_0}, \hat{\omega^{sdid}}) = \text{argmin} \quad l_{unit}(\omega_0, \omega) \tag{6}$$

$$l_{unit}(\omega_0, \omega) = \sum_{t=1}^{T_{pre}} \left(\omega_0 + \sum_{i=1}^{N_{co}} \omega_i Y_{it} - \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} Y_{it} \right)^2 + \zeta^2 T_{pre} ||\omega||_2^2 \tag{7}$$

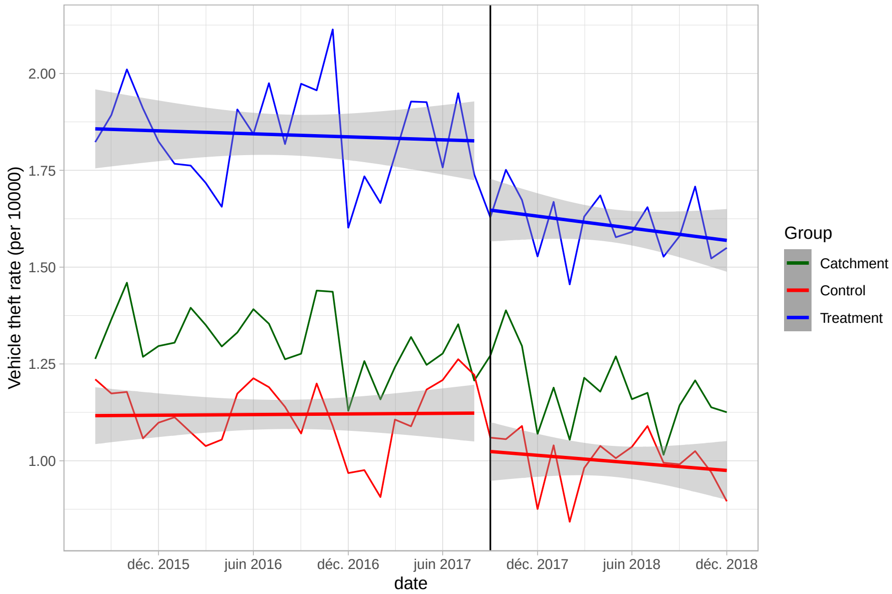


Figure 8: Vehicle theft rates per month from January 2016 to January 2019

Note: this plot represents the trend of vehicle theft rate for the treatment, catchment and control group. Four trend lines are fitted, during the pre-treatment and during the treatment. Two are fitted for the treatment group and two for the control group.

By finding an appropriate intercept for the geographic units ω_0 and adapted individuals weights ω_i, this procedure allow to minimize the quadratic error between the control group outcomes in the pre-treatment period and the treatment group outcomes still in the pre-treatment period. At the end of the procedure, the synthetic difference-in-difference produce an individual weight for each department.

The second part of the optimization program is a regularization of the norm of the weights : by adding the l2 norm of ω to the objective function, the size of the weights is limited. In this case, the ζ is the parameter defining the importance of the penalisation part. Taking a big ζ will lead to more equal ω_i, as $\omega_i = \omega_j$ is the solution which minimize $||\omega||_2^2$.

Using the Synthetic Difference-in-Difference over the the Synthetic control group method has many advantages. First, the presence of ω_0, the units weights intercept, allow us to find a set of ω which only make the trend in both group parallel, and not identical as with the synthetic control group method. Synthetic Difference-in-Difference also allows to re-weight the time units in order to match the pre-trend across both geographical and time dimensions.

However, in this application the analysis if very sensible to the time periods chosen in the control group, as we already observed a strong seasonal effect. Thus, we choose to not re-weight the time periods, but only the geographical units, which is equivalent to set all time weights $\lambda_t = \frac{1}{T_{pre}}$.

For a more detailed explanation of how the weights are effectively obtained, or the algorithmic procedure, please refer to Arkhangelsky et al., 2021.

3 Results

The main results are reported in tables 1, 2 and 3. The results of our preferred specification, model (4), are summarized in Figure 9. We start with the estimates in the *Gendarmerie* areas then in the entire department (*Gendarmerie* plus *Police* areas).

3.1 Estimates of PAVED effect in *Gendarmerie areas*

The results of the estimates for burglary and vehicle theft in the *Gendarmerie* area are reported in Tables 1 and 2 and differ substantially. Note first that in each table, specification (1) shows the effects of the *treatment*, *post*, and *catchment* indicators separately. If we take Table 1, treated *Gendarmerie* areas have, on average, a significantly higher level of burglary by 28% than control areas. This is consistent with the preliminary analysis of the data in the previous sections. The coefficient associated with the *post* variable indicates that the months following the implementation of PAVED ($\text{post}_t = 1$) are marked by a lower average monthly level of burglary in all areas. This is consistent with the pattern observed in the graph in Figure 2. The deterrent effect of PAVED in *Gendarmerie* areas captured by the coefficient associated with the interaction of (*treatment* $\times$ *post*), is weakly positive but not significantly different from zero regardless of the specification used. Note that the magnitude of the effect is reduced by half when fixed-effects

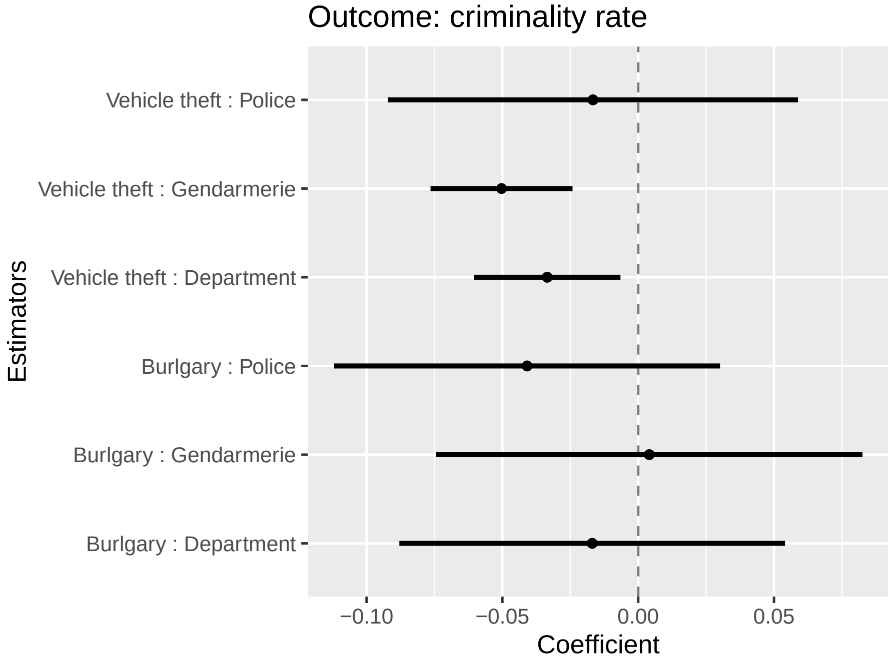


Figure 9: Coefficients from model 4, table 1 to 6

Note: The point indicates the estimated coefficients for the effect of PAVED on the indicated offense in the indicated area. The thick line represents the 95% confidence interval. The estimated model is
$\log(Y_{i,t} + 1) = \alpha_i + \phi_t + \delta_{i,m} + \beta_1\,(\text{treatment}_i \times \text{post}_t) + \beta_2\,(\text{catchment}_i \times \text{post}_t) + \epsilon_{i,t}$. For a detailed table, please refer to table 9

and (department × month) trend are introduced in the specification. The use of the algorithm by the *gendarmes* did not have the expected deterrent effect on burglaries in *Gendarmerie* areas. Reassuringly, the coefficient associated with the (*catchment* × *post*) interaction captures no significant effect regardless of the specification, indicating that PAVED did not generate a displacement of burglaries in catchment areas.

The results in Table 2, reporting the effect of PAVED on vehicle theft in the *Gendarmerie* areas, differ significantly from the results for burglary, reported in Table 1. Consistently with the patterns observed above, the level of vehicle theft is significantly higher in the treated *Gendarmerie* areas than in the control areas, and the post-test periods have lower levels of vehicle theft on average than the pre-test periods. However here, the parameter associated with the (*treatment* × *post*)

Table 1: Effects of PAVED on the log(burglary rate+1) in Gendarmerie areas

Dependent Variable:	log(burglary rate +1)	
Model:	(1)	(2)
Variables		
(Intercept)	1.141***	
	(0.0337)	
treatment	0.2846***	
	(0.0872)	
post	-0.0306*	
	(0.0143)	
catchment	0.0928*	
	(0.0484)	
treatment × post	0.0093	0.0041
	(0.0379)	(0.0400)
catchment × post	0.0048	-0.0055
	(0.0312)	(0.0274)
Fixed-effects		
Date		Yes
Department		Yes
Department × Month		Yes
Fit statistics		
Observations	2,668	2,668
R^2	0.08878	0.80538
Within R^2		0.00014

Clustered (region) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

The estimated model (2) is $\log(\text{burglary rate} + 1) = \alpha_i + \phi_t + \delta_{i,m} + \beta_1 (\text{treatment}_i \times \text{post}_t) + \beta_2 (\text{catchment}_i \times \text{post}_t) + \epsilon_{i,t}$. The indices i correspond to the department, t to the time unit (month × year) and m to the month. The parameter $\beta_1 \times 100$ (resp. β_2) indicate the variation of the monthly burglary in percentage in the treated (resp. catchment) department.

interaction is negative and significantly different from zero at the 1% threshold, whatever the specification used. The magnitude of the effect is around 5 percent

Table 2: Effects of PAVED on the log(vehicle theft rate+1) in Gendarmerie areas

Dependent Variable:	log(vehicle theft rate +1)	
Model:	(1)	(2)
Variables		
(Intercept)	0.6986***	
	(0.0413)	
treatment	0.3084***	
	(0.0734)	
post	-0.0429**	
	(0.0152)	
catchment	0.0540	
	(0.0371)	
treatment × post	-0.0461***	-0.0503***
	(0.0138)	(0.0133)
catchment × post	0.0005	-0.0045
	(0.0163)	(0.0177)
Fixed-effects		
Date		Yes
Department		Yes
Department × Month		Yes
Fit statistics		
Observations	2,668	2,668
R^2	0.09599	0.88991
Within R^2		0.00616

Clustered (region) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

The estimated model (2) is $\log(\text{vehicle theft rate} + 1) = \alpha_i + \phi_t + \delta_{i,m} + \beta_1 (\text{treatment}_i \times \text{post}_t) + \beta_2 (\text{catchment}_i \times \text{post}_t) + \epsilon_{i,t}$. The indices i correspond to the department, t to the time unit (month × year) and m to the month. The parameter $\beta_1 \times 100$ (resp. β_2) indicate the variation of the monthly vehicle theft in percentage in the treated (resp. catchment) department.

and increases slightly when fixed-effects and specific monthly trends by departments are added. This results indicates that use of PAVED has thus generated a deterrent effect on vehicle theft in the *Gendarmerie* areas of the treated departments. This effect is identified as a deterrent effect since the estimate of $catchment \times post$ is not significantly different from zero, indicating the absence of displacement in areas close to the treated *Gendarmerie* areas.

3.2 Estimates of PAVED effect in other areas

Given that the *Gendarmerie* areas represent the overwhelming majority of the territory, it is plausible that PAVED generates effects that spread to the *Police* areas of the treated departments and the entire department. Table 9 (Appendix) indicates the coefficients associated with the PAVED effect on the burglary rate at the department level. Reassuringly, and consistent with the results in Table 1, none of the estimates for ($treatment \times post$) are significantly different from zero indicating no deterrent effect of PAVED on the burglary rate at the department level. There is also no significant effect associated with the variable ($catchment \times post$), which is also consistent since there is no decrease in burglary rates in treated departments that would lead to an increase in burglary rate in nearby departments.

The estimated effects of PAVED on the departmental level of vehicle theft are reported in Table 3. In magnitude, the coefficient associated with the ($treatment \times post$) variable remains negative with a value of 2.5 points, which is about twice as small as the value of the same coefficient for the level of vehicle theft in the *Gendarmerie* areas. The significance of the deterrent effect also decreases with effects at the 10% threshold, and there is no detectable effect of displacement in nearby departments. These results seem to suggest that PAVED produces a clear and local deterrent effect in the *Gendarmerie* areas, but that this effect does not spread, or at least only slightly, to the rest of the department which is by definition made up of *Police* areas.

The results of the estimates of the effects on burglary levels in *Police* areas are reported in Table 9. As expected, we observe neither a deterrent effect of PAVED in the *Police* areas of the treated departments nor a displacement effect in the *Police* areas of departments close to the treated departments. Again, these results

Table 3: Effects of PAVED on the log(vehicle theft rate+1) in the whole department

Dependent Variable:	log(vehicle theft rate +1)	
Model:	(1)	(2)
Variables		
(Intercept)	0.8226***	
	(0.0347)	
treatment	0.3326***	
	(0.0805)	
post	-0.0600***	
	(0.0087)	
catchment	0.1103	
	(0.0665)	
treatment × post	-0.0260*	-0.0260*
	(0.0140)	(0.0141)
catchment × post	-0.0050	-0.0050
	(0.0134)	(0.0135)
Fixed-effects		
Date		Yes
Department		Yes
Department × Month		Yes
Fit statistics		
Observations	2,668	2,668
R^2	0.11706	0.87454
Within R^2		0.00133

Clustered (region) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

The estimated model is $\log(\text{vehicle theft rate} + 1) = \alpha_i + \phi_t + \delta_{i,m} + \beta_1 (\text{treatment}_i \times \text{post}_t) + \beta_2 (\text{catchment}_i \times \text{post}_t) + \epsilon_{i,t}$. The indices i correspond to the department, t to the time unit (month × year), and m to the month. The parameter $\beta_1 \times 100$ (resp. β_2) indicates the variation of the monthly vehicle theft in percentage in the treated (resp. catchment) department.

are consistent with the absence of any detectable effect of PAVED in the treated *Gendarmerie* areas and in the entire departments.

The results of the estimates for vehicle theft in the *Police* areas are in Table 9. The deterrent effect observed on the rate of vehicle theft in the *Gendarmerie* areas and in the entire departments is not observed in the *Police* areas. The estimates associated with the *(treatment* $\times$ *post*) variable have a negative sign but are not statistically different from zero. This result confirms the intuition provided by the results in Table 3: PAVED generates a deterrent effect on the rate of vehicle theft only at the local level in the *gendarmerie* areas where it is used.

3.3 Estimates of PAVED effect - Synthetic DiD

In this section, we present the results of the Synthetic Difference-in-Difference estimation of the effect of PAVED on the burglary and vehicle theft rate in the *Gendarmerie* areas. The objective is to verify if by choosing adequately the department weights in the control group in order to match more closely the pre-trend period, we still obtain a similar estimation of the treatment effect. We use the logarithm of both criminality rates to assess the effect of PAVED in percentages of the criminality level and keep the same pre-treatment period (one year, from 08/2016 to 08/2017). In order to not corrupt the control group with potential displacement effect, we exclude the neighboring departments from the estimation, as they could be chosen in the control group and bias the estimation.

As explained in section 2.6, we match the trend of the control and treatment group only using a re-weighting of geographical units, and not time units. The results are presented in table 4.

The results shows estimation close to those obtained with the TWFE model in precedent section. The treatment effect of PAVED on the burglary rate is still positive and non significant. Once again, we find a negative en significant effect of PAVED on the vehicle theft rate, however this effect is only significant at the 10% level. As explained in the methodology section, we set the pre-treatment time units in the control group to count equally, so every weight is equal to 1/13. All the weights must sum up to 1, and not all the department are effectively included in the control group. They range from 0.017 to 0.042 for the burglary rate estimation,

Table 4: Synthetic Difference-in-Difference estimation of the effect of PAVED on gendarmerie area

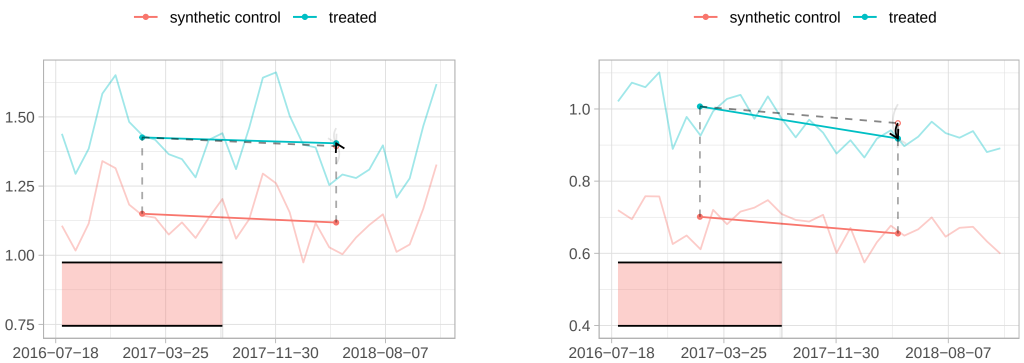
Dep. Var.:	log(Burglary rate +1) (G)	log(Vehicle Theft +1) (G)
Model:	(1)	(2)
Estimator	0.0104	-0.0426*
	(0.0383)	(0.0254)
CI (95%)	[-0.0648;0.0855]	[-0.0923;0.0071]

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

The estimated model is $(\hat{\tau}, \hat{\mu}, \hat{\alpha}, \hat{\beta}) = \text{argmin}\left\{\sum_i^N \sum_t^T (Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2 \hat{\omega}_i^{sdid} \hat{\lambda}_i^{sdid}\right\}$. The indices i correspond to the department, t to the time unit (month $\times$ year). μ is a constant, α_i a geographical unit fixed effect and β_t a time period fixed effect.ω and λ are the individuals geographical units and time periods obtained to fit both control and treatment group pre-trend. The parameter $\tau \times 100$ indicate the variation of the monthly burglary in percentage in the treated department.

with 35 departments in the control group. They range from 0.017 to 0.039 for the vehicle theft rate, with 35 department in the control group.

The pre-trend and treatment effect of PAVED for both estimations are represented in figure 10.



(a) Burglary rate (gendarmerie)

(b) : Vehicle theft rate (gendarmerie)

Figure 10: Synthetic DiD plot

The constant red line under the pre-treatment period indicate that both estimation give equal importance to every periods in the pre-treatment period. As explained, the both trends have to be parallel, and not similar, which is shown here.

3.4 Direct and indirect effect of PAVED

In the precedent sections, we have divided the area of interest into two parts: *Gendarmerie* and *Police* area, and estimated the direct deterrent effect of PAVED on each area and then on both. However, in order to disentangle more precisely the potential deterrence or displacement effect of PAVED in neighboring departments. We use the same two-way fixed effects setting and use the same treatment group definition, but we split the catchment area and use simultaneously both *Gendarmerie* and *Police* area in the same database.

We use three different catchment areas in order to study more finely the displacement or deterrent effect of PAVED on neighboring areas. A neighboring department is defined as sharing a common geographic border with a treatment group department. Thus, the three defined catchment areas are :

- catchment 1 : *Police* area in the treatment departments
- catchment 2 : *Gendarmerie* are in neighboring departments
- catchment 3 : *Police* area in neighboring departments

The control group is the *Gendarmerie* and *Police* areas neither in the treatment nor the neighboring department group (called catchment areas). We include both jurisdictions in the control group to have an adapted counterfactual for the entirety of the catchment area. While the treatment group and catchment 2 are *Gendarmerie* areas, and so are suitable to be compared to only *Gendarmerie* area, catchment 1 and catchment 3 groups are *Police* area and are best to be compared to *Police* area also. We end up with 11 geographical units in the treated group, 11 in catchment 1, 76 in catchment 2, 76 in catchment 3 and, 86 geographical units in the control group. The pre-treatment period is from 08/2016 to 08/2017, while the treatment period spans from 09/2017 to 12/2018.

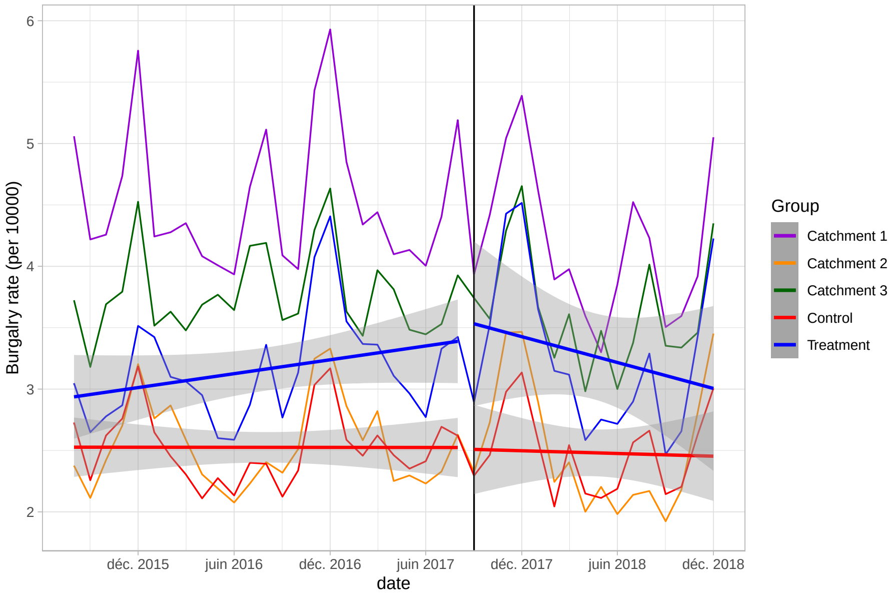


Figure 11: Trend of all groups for burglary rate

The trend for both criminality indicators a plotted in figure 11 and 12. By using *Police* and *Gendarmerie* area in the control group, it appears that the parallel trend for the burglary rate is less convincing that in the last sections. The parallel trend for control and treatment groups regarding the vehicle theft rate is however convincing and comfort the use of this specification and database to identify the effect of PAVED on treated areas. Observing the regression lines, we can distinguish a bigger drop in the trend of the treatment group than the control group, which has to be verified using the augmented specification of the TWFE model.

The two-way fixed effects model is similar except for two indicators: the geographic indicator is at the sub-department level (*Gendarmerie* or *Police* area in a given department). We also include a jurisdiction fixed effect to control for the heterogeneity between *Gendarmerie* and *Police* areas. The date and department $\times$ month (seasonal effect of each department) are the same as in the last precedent section.

We estimate the following model for the both measurement of criminality, the burglary and vehicle theft rate, log transformed :

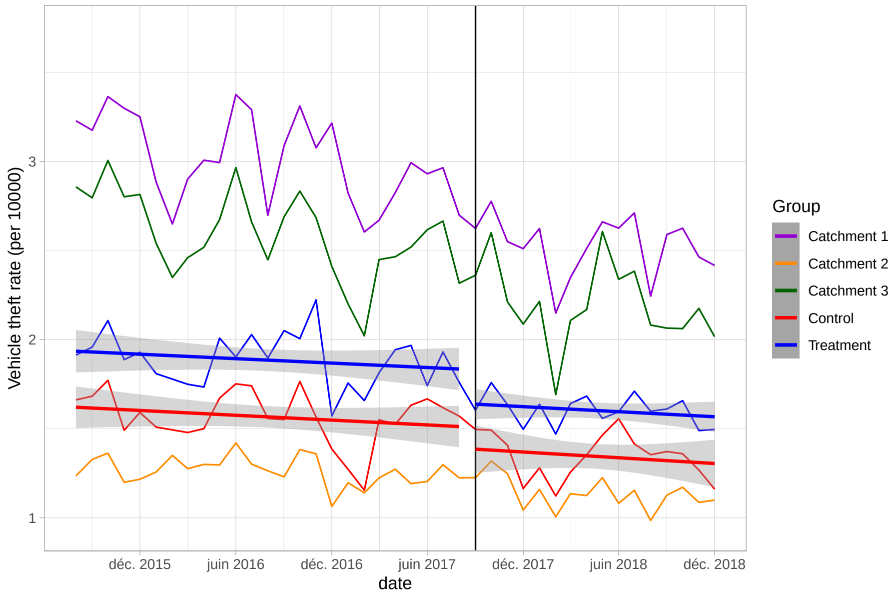


Figure 12: Burglary rates by region (2018)

$$\log\left(\frac{c_{kt}}{pop_{kt} \times \frac{1}{10{,}000}} + 1\right) = \delta_{i,m} + \eta_j + \gamma_t + \beta_1(treatment \times post) + \beta_2(catchment_1 \times post) +$$
$$\beta_3(catchment_2 \times post) + \beta_4(catchment_3 \times post) + \epsilon_{jt}$$

With k the type of crime, i the department, m the month, j the sub-department area, and t the date (month $\times$ year). $\delta_{i,m}$ is the seasonal effect of each department, η_j the geographical fixed effect, and γ_t the time unit fixed effect. We implement first a regional then sub-department fixed effect.The results are presented in table 5 for the burglary rate and in 6 for the vehicle theft rate.

We produce the two models by adding an increasing precision in the fixed effect at each step. We succeed to reproduce similar results as in table 1. The estimated parameters associated with *(treatment* and *post*) indicate that the treatment area has a higher burglary rate average than the control group and that the trend in this rate is decreasing in the short run, at least in the database period.

The associated effects of *(catchment 1 : catchment 3*) are indicative of the structure of burglary in France. We can observe that the first catchment area exhibit the highest burglary rate of all defined areas, especially higher than the *treatment* parameter. This result reinforces the fact that the tested department is more prone to criminality, and the urban area is more exposed on average than the rural part. This criminal structure is similar in the catchment areas: the *Police* areas of neighboring department (*catchment 3*) have on average a higher burglary rate than the *Gendarmerie* areas (*catchment 2*).

We find a similar structure of the effect as in table 1 with a weak positive effect. The difference between both estimations is explained by the introduction of a smaller scale geographical fixed effect *sub department*, the fixed effect *juridiction*, and the variation of the control group which includes both jurisdictions. The estimated parameters of (*treatment* $\times$ *post*) completely shrink when the last layer of fixed effect is added, confirming the absence of effects of PAVED on the burglary rate.

As expected, none of the three catchment areas show a significant effect, as the criminality rate in treated areas is not impacted.

Model 1 replicates the structure of the effect obtained in table 1 with slight variation. The intercept is now higher as the control group includes *Police* areas, which have a higher average criminality rate. Reassuringly, *treatment* and *Intercept* still sum up to 1.007. We find again the same effect structure, with the *Police* area of treated departments exhibiting the highest vehicle theft rate, followed by *Police* in neighboring departments, *Gendarmerie* in treated departments, and *Gendarmerie* in neighboring department. The decreasing temporal trend of vehicle theft is still observable with a negative estimated parameter for *post*.

The specification 3.4 succeeds to capture the deterrent effect of PAVED on vehicle theft, as we can identify a negative, significant, and persistent treatment effect ranging from -2.2% to -3% from model 1 to model 2. This result is very similar to the one from table 2, which is encouraging regarding the stability of this effect.

The potential displacement or deterrent effect is identified in three potential catchment areas. It is important to understand that these two effects compete with each other, as if we observe an increase in the criminality rate of neighboring area, this will indicate a displacement effect: criminals adapted their behavior following the increase in *Gendarmerie* patrols productivity and went to a different geographic area. However, if we observe a decrease in the neighboring area, the deterrent effect of PAVED on the treated area also produces spillover in the surrounding areas. In this estimation, we identify a slight displacement effect from the treated area to the neighboring *Gendarmerie* area (*catchment 2*). This effect is significant at the 10% level and identified from model 1. However, once we add fixed-effects, we cannot identify this effect anymore. Thus, it doesn't allow to conclude on a potential displacement effect of criminality in the neighboring areas following the positive productivity shock that PAVED had on patrol units.

On the direct effect of PAVED on the vehicle theft rate, we can conclude that this effect is ranging from -5% to -3% on average on the vehicle theft rate of treated departments, over the test period. Referring to table 8, this reduction is equivalent to a decrease of 114 to 68 vehicle theft per department and year.

Table 5: Effects of PAVED on burglary rate in treatment and catchment areas

Dependent Variable:	log(burglary rate +1)	
Model:	(1)	(2)
Variables		
(Intercept)	1.213***	
	(0.0469)	
treatment	0.2127**	
	(0.0878)	
post	-0.0240*	
	(0.0115)	
catchment 1 (police in dep_w)	0.4447***	
	(0.1150)	
catchment 2 (gendarmerie in dep_n)	0.0210	
	(0.0601)	
catchment 3 (police in dep_n)	0.2707***	
	(0.0840)	
treatment × post	0.0026	0.0003
	(0.0312)	(0.0338)
post × catchment 1	-0.0349	-0.0372
	(0.0364)	(0.0380)
post × catchment 2	-0.0018	-0.0084
	(0.0276)	(0.0247)
post × catchment 3	-0.0081	-0.0146
	(0.0252)	(0.0253)
Fixed-effects		
Date		Yes
Sub department		Yes
Department × Month		Yes
Fit statistics		
Observations	5,336	5,336
R^2	0.13928	0.75801
Within R^2		0.00068

Clustered (region) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

The estimated model is $\log(\text{burglary rate}{+}1) = \delta_{i,m} + \eta_j + \gamma_t + \zeta_j + \beta_1(treatment \times post) + \beta_2(catchment_1 \times post) + \beta_3(catchment_2 \times post) + \beta_4(catchment_3 \times post) + \epsilon_{jt}$. The indices i correspond to the department, j to the sub-department, t to the time unit (month × year), and m to the month. The parameter $\beta_1 \times 100$ (resp. β_2, β_3 and β_4) indicates the variation of the monthly burglary in percentage in the treated department.

Table 6: Effects of PAVED on vehicle theft rate in treatment and catchment areas

Dependent Variable:	log(vehicle theft rate +1)	
Model:	(1)	(2)
Variables		
(Intercept)	0.8700***	
	(0.0326)	
treatment	0.1370*	
	(0.0720)	
post	-0.0669***	
	(0.0076)	
catchment 1 (police in dep_w)	0.4595***	
	(0.0811)	
catchment 2 (gendarmerie in dep_n)	-0.1174**	
	(0.0460)	
catchment 3 (police in dep_n)	0.2843***	
	(0.0670)	
treatment × post	-0.0221*	-0.0302**
	(0.0119)	(0.0114)
post × catchment 1	-0.0287	-0.0368
	(0.0310)	(0.0314)
post × catchment 2	0.0245*	0.0209
	(0.0124)	(0.0141)
post × catchment 3	-0.0213	-0.0249
	(0.0151)	(0.0171)
Fixed-effects		
Date		Yes
Sub department		Yes
Department × Month		Yes
Fit statistics		
Observations	5,336	5,336
R^2	0.19669	0.83476
Within R^2		0.00356

Clustered (region) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

The estimated model is $\log(\text{vehicle theft rate} + 1) = \delta_{i,m}+\eta_j+\gamma_t+\zeta_j+\beta_1(treatment\times post)+\beta_2(catchment_1\times post) + \beta_3(catchment_2\times post)+\beta_4(catchment_3\times post)+\epsilon_{jt}$. The indices i correspond to the department, j to the sub-department, t to the time unit (month × year), and m to the month. The parameter $\beta_1 \times 100$ (resp. β_2, β_3, β_4) indicates the variation of the monthly vehicle theft in percentage in the treated (resp. catchment) department.

4 Conclusion

As predictive methods are increasingly used by law enforcement agencies to predict criminality and improve their response, public evaluations of these systems become crucial. In France in particular, little information is made public on either the implementation or the effects of this type of device. In this paper, we provide the first evaluation of PAVED, a predictive algorithm used by the *Gendarmerie* in 11 departments a little over a year, which jurisdiction covers 95% of the territory and 50% of the population. Based on public information and data, we test whether the algorithm fulfills the stated goal of crime reduction through deterrence and not through displacement.

Our panel data on burglaries and vehicle thefts observed monthly in all the *Gendarmerie* and *Police* areas allow us to implement several complementary estimation methods. We first rely on an augmented Two Way Fixed Effect model including a seasonal-department fixed effect, which we applied on a sub-segment of France and then on the whole country. Then, in order to improve the credibility of the parallel pre-trends hypothesis, mainly for burglaries, we use the synthetic diff in diff. This method consists of constructing a control group more similar to the treatment group to make the linear pre-trends of the two groups also more similar. We get much more convincing trends for burglaries.

According to these different estimation methods, we find an effect of PAVED ranging from -5% to -3% on vehicle theft in the tested *Gendarmerie* areas. This reduction corresponds to an average of around 100 vehicle thefts avoided per department per year. Conversely, we do not capture any reduction effect of PAVED on burglaries, regardless of the areas considered and regardless of the estimation methods employed. The effects of PAVED in neighboring areas are mostly insignificant. We capture a displacement effect of vehicle theft in *Gendarmerie* areas in neighboring departments, but this effect is not robust across specifications. These results suggest that the reduction effect on vehicle theft is achieved through a deterrent effect on criminals rather than a displacement effect in nearby areas.

We see two main explanations for these results. First, it is plausible that the effectiveness of *gendarme*'s patrols in areas indicated by PAVED is greater for vehicle theft than for burglary. Most vehicles are parked on the street and very regular street patrols may significantly increase criminals' perception of the likelihood of being caught stealing a vehicle. This increase in the perceived probability of arrest potentially significantly increases deterrence. On the other hand, burglaries take place out of sight, inside the houses, and criminals may feel less exposed due to the increased patrols. Second, it is plausible that the gendarmes' initial level of knowledge regarding burglary and vehicle theft hotspots differs. Our results suggest that PAVED does not add value to *gendarmes*' initial knowledge for burglary, whereas it does for vehicle theft.

We hope this article show how empirical evaluations of these applications are needed to understand the effect of predictive systems on criminality. However, this paper only investigates the "benefit" part while these systems come at different costs: design and maintenance costs but also potentially social costs if the algorithm is not properly trained. If the data feeding the algorithm is biased, the algorithm can replicate and amplify that bias. If the bias in question stem from discriminatory law enforcement behavior, then the algorithm potentially generate significant social costs.

Although the French law provides a framework for the openness of public algorithms and rules regarding public use of these systems[15], this framework doesn't apply for law enforcement agency. Thus, it is not legally required for the *Gendarmerie* to publish general information regarding the algorithm, to mention when it is used, and to provide individual information regarding person's treatment of the algorithm. These information would be essential for future research and evaluation of predictive policy algorithms.

[15] Code des relations entre le public et l'administration (CRPA): https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000033205535

References

Apel, Robert (2013). “Sanctions, Perceptions, and Crime: Implications for Criminal Deterrence”. In: *Journal of Quantative Criminology* 29, pp. 67–101.

Apel, Robert and Daniel S. Nagin (2011). “General Deterrence”. In: *The Oxford Handbook of Crime and Criminal Justice*.

Arkhangelsky, Dmitry et al. (Dec. 1, 2021). “Synthetic Difference-in-Differences”. In: *American Economic Review* 111.12, pp. 4088–4118. ISSN: 0002-8282. DOI: 10.1257/aer.20190159. URL: https://pubs.aeaweb.org/doi/10.1257/aer.20190159 (visited on 05/24/2022).

Becker, Gary S (1968). “Crime and Punishment: An Economic Approach”. en. In: *Journal of Political Economy* 76.2, pp. 169–217.

Braga, Anthony A. (2008). *Police Enforcement Strategies to Prevent Crime in Hot Spot Areas*. Tech. rep. Washington, D.C.: U.S. Department of Justice Office of Community Oriented Policing Services.

Chalfin, Aaron and Justin McCrary (Mar. 2017). “Criminal Deterrence: A Review of the Literature”. In: *Journal of Economic Literature* 55.1, pp. 5–48. DOI: 10.1257/jel.20141147. URL: https://www.aeaweb.org/articles?id=10.1257/jel.20141147.

Durlauf, Steven N., Salvador Navarro, and David A. Rivers (Oct. 2010). “Understanding aggregate crime regressions”. In: *Journal of Econometrics* 158.2, pp. 306–317. ISSN: 03044076. DOI: 10.1016/j.jeconom.2010.01.003. URL: https://linkinghub.elsevier.com/retrieve/pii/S0304407610000060 (visited on 05/24/2022).

Germes, Melina (2014). “Cartographies policières : la dimension vernaculaire du contrôle territorial”. In: *EchoGéo* 28, p. 30.

Guerette, Rob T. and Kate J. Bowers (2009). “ASSESSING THE EXTENT OF CRIME DISPLACEMENT AND DIFFUSION OF BENEFITS: A REVIEW OF SITUATIONAL CRIME PREVENTION EVALUATIONS*”. In: *Criminology* 47.4, pp. 1331–1368. DOI: https://doi.org/10.1111/j.1745-9125.2009.00177.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1745-9125.2009.00177.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1745-9125.2009.00177.x.

Ho, Daniel E., John J. Donohue, and Patrick Leahy (2013). “Empirical Legal Analysis: Assessing the Performance of Legal Institutions”. In: ed. by Yun-Chien Chang. Routledge. Chap. Do Police Reduce Crime? A Reexamination of a Natural Experiment.

Jabri, Ranae (2021). “Algorithmic Policing”. DOI: 10.2139/ssrn.4275083. URL: https://ssrn.com/abstract=4275083.

Levitt, Steven D. (1997). “Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime”. In: *The American Economic Review* 87.3, pp. 270–290. ISSN: 00028282. URL: http://www.jstor.org/stable/2951346 (visited on 07/07/2022).

Lochner, Lance (2004). “Education, Work, and Crime: A Human Capital Approach”. In: *International Economic Review* 45.3, pp. 811–843. ISSN: 00206598, 14682354. URL: http://www.jstor.org/stable/3663638 (visited on 06/29/2022).

Mastrobuoni, Giovanni (Nov. 1, 2020). “Crime is Terribly Revealing: Information Technology and Police Productivity”. In: *The Review of Economic Studies* 87.6, pp. 2727–2753. ISSN: 0034-6527, 1467-937X. DOI: 10.1093/restud/rdaa009. URL: https://academic.oup.com/restud/article/87/6/2727/5799090 (visited on 05/24/2022).

Nano, Edlira and Félix Tréguer (Sept. 2022). *Place-Oriented Predictive Policing in France*. Research Report. Technopolice Research Collective. URL: https://hal.archives-ouvertes.fr/hal-03783592.

Polinsky, A. Mitchell and Steven Shavell (2000). “The Economic Theory of Public Enforcement of Law”. In: *Journal of Economic Literature* XXXVIII, pp. 45–76.

Weisburd, David L. et al. (2006). “Does Crime Just Move Around the Corner? A Controlled Study of Spatial Displacement and Diffusion of Crime Control Benefits”. In: *Criminology* 44.3, pp. 549–591.

Appendix

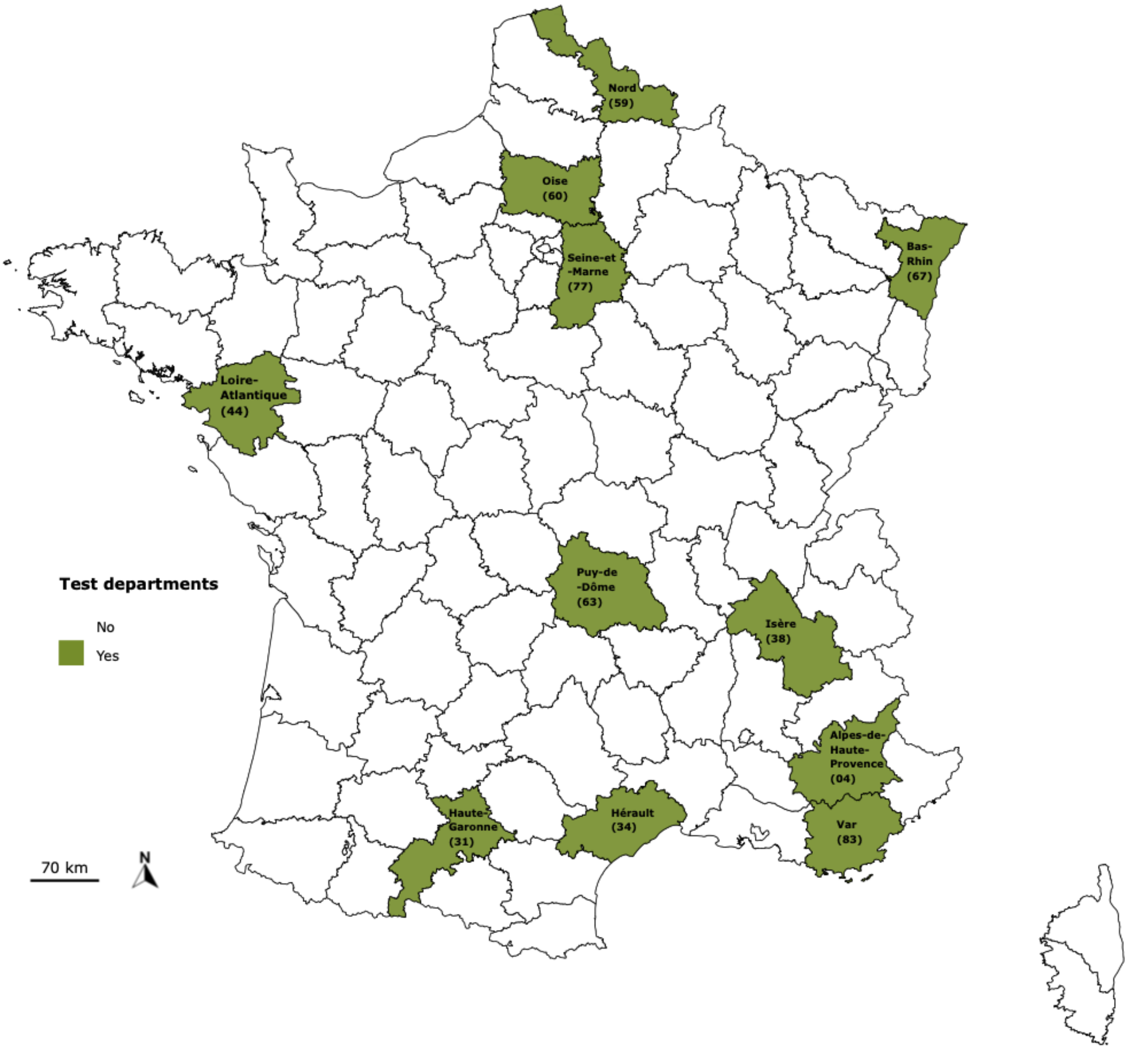


Figure 13: Departments tested over the period September 2017-March 2019.

4.1 Descriptive statistics

Geographical unit	Burglary rate	B. rate (Police)	B. rate (Gendarmerie)	Pop.
04	3.72	3.48	3.80	164068
31	4.62	6.32	3.34	1380672
34	3.90	4.48	3.46	1159220
38	4.14	5.16	3.73	1263563
44	3.71	5.16	2.76	1412502
59	3.17	3.45	2.37	2606234
60	2.58	2.30	2.65	827153
63	2.87	3.54	2.35	659048
67	1.75	2.21	1.45	1133552
77	3.35	3.41	3.19	1412516
83	4.64	4.54	4.75	1067697
Treatment group	3.50	4.00	3.08	1189657
Control group	2.53	3.09	2.21	540099
France (w.o Paris)	2.64	3.20	2.32	617763

Table 7: Burglary rate for all treated departments, treatment group, control group and France in 2018 (for 10.000 inhabitants)

Geographical unit	Vehicle Theft r.	VT. r. (Police)	VT. r. (Gendarmerie)	Pop.
04	1.36	1.78	1.22	164068
31	1.61	2.24	1.13	1380672
34	2.11	2.73	1.64	1159220
38	2.15	2.77	1.91	1263563
44	2.26	3.30	1.57	1412502
59	2.47	2.84	1.43	2606234
60	3.11	3.86	2.89	827153
63	1.63	2.14	1.23	659048
67	0.66	1.08	0.38	1133552
77	2.49	2.47	2.55	1412516
83	1.92	2.20	1.59	1067697
Treatment group	1.98	2.49	1.60	1189657
Control group	1.34	1.90	1.04	540099
France (w.o Paris)	1.42	1.97	1.11	617763

Table 8: Vehicle theft rate for all treated departments, treatment group, control group and France in 2018 (for 10.000 inhabitants)

4.2 Gendarmerie, Police and whole department tables for the burglary and vehicle theft rate

Table 9: Recap of all the model 4 results, for each indicators and geographic areas

Dependent Variables:	1	2	3	4	5	6
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
treatment × post	0.0041	-0.0503***	-0.0409	-0.0166	-0.0169	-0.0335**
	(0.0400)	(0.0133)	(0.0362)	(0.0385)	(0.0362)	(0.0138)
catchment × post	-0.0055	-0.0045	-0.0174	0.0005	-0.0058	-0.0092
	(0.0274)	(0.0177)	(0.0219)	(0.0182)	(0.0230)	(0.0147)
Fixed-effects						
date	Yes	Yes	Yes	Yes	Yes	Yes
dep	Yes	Yes	Yes	Yes	Yes	Yes
dep × month	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	2,668	2,668	2,668	2,668	2,668	2,668
R^2	0.80538	0.88991	0.79137	0.81010	0.86420	0.92601
Within R^2	0.00014	0.00616	0.00128	0.00029	0.00051	0.00360

Clustered (regionn) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

The model estimated is :

$$\log(Y_{i,t}+1) = \alpha_i+\phi_t+\delta_{i,m}+\beta_1\,(\text{Treatment}_i \times \text{Post}_t)+\beta_2\,(\text{Catchment}_i \times \text{Post}_t)+\epsilon_{i,t} \tag{8}$$

This is a detailed recap of figure 9. We use the same range of fixed effect as in model 4 : department, date and department × month.

Dependent variables :

- 1 : log(burglary rate +1) in gendarmerie area
- 2 : log(vehicle theft rate +1) in gendarmerie area
- 3 : log(burglary rate +1) in police area
- 4 : log(vehicle theft rate +1) in police area
- 5 : log(burglary rate +1) in whole department area
- 6 : log(vehicle theft rate +1) in whole department area

4.3 Placebo test

In order to test if the identified effect is not generated by a random noise in the data, we proceed to three placebo tests. We randomly select 11 department in the control groupe and test for the presence of an effect in these groups. The selected placebo groups are :

- placebo 1 : 55, 64, 06, 36, 25, 54, 78, 19, 68, 62, 66
- placebo 2 : 72, 37, 81, 43, 58, 71, 50, 89, 91, 19, 23
- placebo 3 : 2A, 08, 68, 89, 46, 18, 80, 91, 86, 2B, 02

Table 10: Effects of PAVED on the log(vehicle theft rate+1) in Gendarmerie areas, placebo test

Dependent Variable:	log(vehicle theft rate +1)		
Model:	(1)	(2)	(3)
Variables			
placebo 1 × post	-0.0091		
	(0.0222)		
placebo 2 × post		0.0015	
		(0.0335)	
placebo 3 × post			-0.0181
			(0.0224)
Fixed-effects			
date	Yes	Yes	Yes
department	Yes	Yes	Yes
department ×mois	Yes	Yes	Yes
Fit statistics			
Observations	2,349	2,349	2,349
R^2	0.87380	0.87377	0.87389
Within R^2	0.00023	6.78×10^{-6}	0.00093

Clustered (regionn) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

4.4 Longer pre-treatment period

In this article, we choose to use a pre-treatment time period of one year. As this constraint can appear restrictive, we run the same estimation as in table 2, but by using two years as the pre-treatment period (from 09/2015 to 08/2017). The estimated parameters of the effect of PAVED is in the same range as our first estimation, with a similar significance level.

Table 11: Effects of PAVED on the log(vehicle theft rate+1) in Gendarmerie areas, two years pre-treatment period

Dependent Variable:	log(vehicle theft rate +1)		
Model:	(1)	(2)	(3)
Variables			
(Intercept)	0.7003***		
	(0.0475)		
treatment	0.3133***		
	(0.0786)		
post	-0.0446**		
	(0.0189)		
catchment	0.0629		
	(0.0406)		
treatment × post	-0.0510**	-0.0510**	-0.0543**
	(0.0222)	(0.0223)	(0.0225)
catchment × post	-0.0083	-0.0083	-0.0098
	(0.0205)	(0.0206)	(0.0213)
Fixed-effects			
date		Yes	Yes
department		Yes	Yes
department × month			Yes
Fit statistics			
Observations	3,680	3,680	3,680
R^2	0.09702	0.79923	0.86341
Within R^2		0.00322	0.00518

Clustered (regionn) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

4.5 High and Low criminality treatment groups

4.5.1 High criminality

Table 12: Effects of PAVED on the log(vehicle theft rate+1) in Gendarmerie areas, high criminality treatment group subsample

Dependent Variable:	log(vehicle theft rate +1)		
Model:	(1)	(2)	(3)
Variables			
(Intercept)	0.6986***		
	(0.0413)		
treatment_{high}	0.5244***		
	(0.0799)		
post	-0.0429**		
	(0.0152)		
catchment	0.0540		
	(0.0371)		
$\text{treatment}_{high} \times$ post	-0.0589*	-0.0589*	-0.0634**
	(0.0314)	(0.0315)	(0.0284)
post × catchment	0.0005	0.0005	-0.0045
	(0.0163)	(0.0164)	(0.0177)
Fixed-effects			
date		Yes	Yes
department		Yes	Yes
department × month			Yes
Fit statistics			
Observations	2,494	2,494	2,494
R^2	0.15049	0.80646	0.89147
Within R^2		0.00279	0.00520

Clustered (regionn) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 13

4.5.2 Low criminality

Table 14: Effects of PAVED on the log(vehicle theft rate+1) in Gendarmerie areas, low criminality treatment group subsample

Dependent Variable:	log(vehicle theft rate +1)		
Model:	(1)	(2)	(3)
Variables			
(Intercept)	0.6986***		
	(0.0413)		
treatment_{low}	0.1284		
	(0.1080)		
post	-0.0429**		
	(0.0152)		
catchment	0.0540		
	(0.0371)		
treatment_{low} × post	-0.0354	-0.0354	-0.0394
	(0.0354)	(0.0356)	(0.0364)
catchment × post	0.0005	0.0005	-0.0045
	(0.0163)	(0.0164)	(0.0177)
Fixed-effects			
date		Yes	Yes
department		Yes	Yes
department × month			Yes
Fit statistics			
Observations	2,523	2,523	2,523
R^2	0.02150	0.78008	0.87340
Within R^2		0.00119	0.00227

Clustered (regionn) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

4.6 Synthetic Difference-in-Difference weights

Weights for R1		Weights for R2	
department	weight	department	weight
47	0.042	72	0.039
74	0.037	41	0.039
25	0.032	78	0.035
40	0.032	21	0.033
21	0.032	91	0.032
28	0.032	52	0.032
71	0.031	37	0.032
78	0.030	25	0.030
52	0.030	90	0.030
55	0.029	36	0.029
37	0.028	48	0.029
66	0.028	55	0.029
54	0.028	54	0.028
70	0.028	28	0.027
72	0.027	79	0.026
29	0.027	8	0.025
8	0.026	50	0.025
14	0.026	46	0.025
90	0.026	66	0.025
39	0.025	86	0.024
2A	0.025	14	0.024
7	0.024	74	0.024
43	0.024	39	0.022
53	0.024	87	0.022
24	0.024	29	0.021
16	0.023	58	0.021
33	0.022	61	0.021
22	0.022	53	0.021
61	0.021	64	0.020
2B	0.021	16	0.020
91	0.019	24	0.019
50	0.019	70	0.019
86	0.018	43	0.019
46	0.018	17	0.018
17	0.017	40	0.017