Pre-Analysis Plan for "Impact evaluation of a new homicide investigation policy in Bogotá D.C., Colombia"

Daniela Collazos   Leopoldo Fergusson   Miguel E. La Rota   Daniel Mejía   Daniel Ortega*

March 13, 2017

Contents

_____
*Collazos: Universidad de los Andes, Economics, d.collazos148@uniandes.edu.co; Fergusson: Universidad de los Andes, Economics, lfergusson@uniandes.edu.co; La Rota: Fiscalía General de la Nación, Law and Economics, miguel.larota@fiscalia.gov.co; Mejía: Universidad de los Andes, Economics, dmejia@uniandes.edu.co; Ortega: Latin America Development Bank - CAF, dortega@caf.com.

1 Introduction

1.1 Motivation

Homicide investigation studies focus on identifying and analyzing best practices to solve homicide cases and increase the probability that the offender is prosecuted for the crime. The empirical academic research on this topic is scarce, and to the best of our knowledge there is no rigorous evaluation of the effectiveness of alternative investigation strategies. We contribute filling this gap with a randomized controlled trial evaluating a change in the homicide investigation process in Bogotá D.C., Colombia. The main objective of the new policy is to improve the investigation process and its judicial outcomes.

The context is ideal bearing in mind the poor performance of homicide investigation, evident in very low indictment rates (i.e., the percentage of criminal cases where a person is charged with committing the crime). In 2015, Colombia's homicide rate was 24 per 100,000 people, its indictment rate 21.5%, and its conviction rate (percentage of cases with a conviction, out of those with indictment) was 62%. For Bogotá, the homicide rate was 18 per 100,000 people, its indictment rate 22.4%, and its conviction rate 74.6%.

A similar measure used in the literature is the homicide clearance rate, or the percentage of total cases in a year that are solved in that same time period. According to the FBI's uniform crime reports (UCR) an offense is solved or cleared if "at least one person is arrested, charged with the commission of the offense, and turned over to the court for prosecution"[1]. In 2015, 61.5% of murder offenses were cleared by arrest or exceptional means in the United States. When contrasting with the Colombian figures of mere one out of five simply indicted, it is clear that an understanding of how to improve judicial efficiency is key.

1.2 Background: the homicide investigation process in Colombia

In Colombia, two national divisions are responsible of homicide investigation: The Sectional Division of Crime Investigation of the National Police (SIJIN, for its Spanish acronym) and the Technical Division of Crime Investigation of the Attorney General Office (CTI). The homicide investigation policy we evaluate was developed and implemented only within the latter. The Automatic Dispatch Center of the National Police (CAD) responds to an initial complaint—by civilians or police authorities at the crime scene— and informs a Satellite Unit (*Unidad Satélite*). The unit then assigns the case to either the SIJIN or the CTI. For CTI cases, when the initial complaint is made, a District Attorney (DA) on duty closest to the Immediate Reaction Unit (URI) is also assigned to the investigation.

The criminal procedure for homicides in Bogota D.C. has three major stages. The first stage is the initial investigation (*Investigación*) of the crime, in which physical evidence is collected in order to identify and arrest the offender. The second is the pre-trial stage (*Indagación*). It encompasses the indictment (*Imputación*) (i.e. the formal accusation against a person for the criminal offense, done in Colombia only before a Judge) and the gathering, if necessary, of corroborating evidence

that proves beyond reasonable doubt the culpability of the suspect. The final stage is the trial, starting with the arraignment hearing (*Audiencia de formulación de acusación*) in which the accused pleads guilty or not guilty and the evidence is outlined to determine if it is enough to proceed to trial.

1. Initial investigation (*Investigación)*

Once the CTI is assigned to investigate the case, it dispatches a team of crime scene investigators (CSIs) and two detectives. The CSIs document the scene, collect material evidence and prepare the body to transport it to the National Institute of Legal Medicine and Forensic Sciences (INML). At the INML the body is processed for external evidence and a forensic pathologist performs the autopsy. The detectives interview witnesses and follow leads, and execute any orders of the URI's DA leading the investigation. After documenting the evidence, they write detailed reports of their activities at the crime scene: the CSI leader writes the record of technical examination of the corpse (*Acta de inspección técnica a cadáver*) and the leading detective drafts the Executive Report (*Informe ejecutivo)*.

The URI's DA decides—based on these reports—if further enquiry is needed, and if so, instructs detectives to keep investigating. Then, he or she (or other DA within the URI) prepares the methodological program (*Programa metodológico*), summarizing the facts and proposing the criminal hypotheses, and sends it to a DA of the Life Unit (*Unidad de Vida*), to whom the case is transferred. However, at this point further transfers may occur to other DAs within the Life Unit. Life Unit DAs decide whether to continue or not with the criminal procedure—in the event a suspect was identified during the initial investigation— thus initiating the pre-trial stage, or to classify the investigation as an unsolved case if there are no active leads that detectives can pursue.

2. Pre-trial (*Indagación)*

The pre-trial stage begins when a suspect is formally charged. When the offender is arrested, the DA must bring the detainee before a judge of guarantees (*juez de garantías*) within 36 hours to ensure the legality of the procedure and to determine if the person should be released on bail or held in custody. If the person is arrested *in flagrante* the initial investigation and the indictment occur simultaneously. Colombia's Penal Code defines *in flagrante* as the situation when a person is (i) caught and apprehended during the commission of the crime, or (ii) the person is caught or identified during the commission of the crime and apprehended immediately after by hot pursuit or calls for help by witnesses or (iii) when the person is caught and captured with objects indicating probable culpability of having committed the crime. The DA has 30 days after the indictment to continue with the next step of the criminal procedure, the arraignment, or to ask the judge for a dismissal.

During this stage, detectives can and should continue searching for evidence to determine beyond a doubt whether the suspect committed the crime.

3. Trial (*Juicio*)

The arraignment (*Formulación de acusación)* is the first step of this stage. It is a procedure whereby the accused is brought before the judge to plea to the criminal charge in the indictment. The indictment bill (*Escrito de acusación)* is read to the defendant so he or she understands what the

charges are, what his or her constitutional rights are, and then he or she is asked by the judge to plead guilty or not guilty. When pleading guilty, the case proceeds to sentencing. Otherwise, the judge holds a preparatory hearing (*Audiencia preparatoria)* to evaluate if there is enough evidence that the defendant committed the crime and if the case should continue to trial. In Colombia, trials are conducted before a judge without a jury.

1.3 Diagnosis of the homicide investigation process

Many factors can affect the results of a homicide investigation process and it is important to identify problems and best practices in order to improve the effectiveness of the investigations. Brookman & Innes (2013) identify four definitions of investigative success, of which we highlight two: outcome success, which concerns the identification, prosecution and conviction of the suspects, and, procedural success, which is about maintaining the integrity and quality of the investigation. These definitions of success enclose the main objectives of the intervention: improving the initial investigation process to identify the perpetrator and obtaining sufficient evidence to establish beyond reasonable doubt his or her culpability.

Several aspects determine the success of an investigation, and the final outcome could be sensitive to factors related with the circumstances of the crime and the capacity of the homicide investigation unit and the judicial system. For example, factors such as the lack of resources, a poor working relationship between DAs and investigators, and poor procedures for processing and analyzing evidence can negatively affect the murder clearance rate (Cronin et al, 2007). Instead, best practices such as formal training of homicide detectives and the use of sophisticated analytical tools altogether improve the solvability of homicides (Keel, Jarvis & Muirhead, 2009). Also, a cooperative relationship between DAs and detectives is related with better clearance rates and with successful prosecutions.

Bernal & La Rota (2014) conducted a diagnosis of the current homicide investigation process in Colombia and identified, among others, two main problems: (1) there are multiple officers acting at different stages under a diffuse leadership and (2) there are delays generated by successive reallocations between different prosecutors of the URI and the Life Unit. This breaks the link between the three stages of the criminal procedure. For the authors, this disruption is caused by the lack of leadership taking responsibility for the case from the beginning.

Based on this analysis, the Attorney's office concluded that the current homicide investigation structure needed to be restructured into a more efficient one, avoiding the loss of information between public servants and more effectively using time, human and physical resources. The Office for Public Policy Affairs of the Attorney General developed a new structure for the murder investigation process to overcome these problems. We describe this new policy in the next section.

2 Intervention

2.1 Description of the intervention

In our intervention, a fraction of homicide cases under the CTI´s responsibility follow a new investigative procedure. This new procedure seeks to overcome the disconnection between the initial investigation, the pre-trial, and the trial. The main change is structuring teams that investigate murder cases together, from the initial assignment or urgent acts (*actos urgentes)* until the

arraignment. Each case is led by the DA and is investigated by a CSI team (consisting of three to four people) and two detectives. From now on, we will thus refer to a homicide squad or group as a team composed by one DA, one CSI team and two detectives. Each unit is simultaneously responsible for several cases.

The introduction of homicide squads changes the existing process of homicide investigation in the following main dimensions.

First, from the moment the CSI and the detectives are dispatched to the crime scene they know who the DA in charge of the investigation is. This allows them to communicate, receive direct orders, and clarify possible questions.

Second, the same DA is responsible from the urgent acts until the arraignment. This eliminates the URI's involvement and possible transfer to a different DA. The only exceptions are when the case is transferred to a DA specialized on homicides with specific characteristics: juvenile offenders, or when the forensic pathologist report is required to determine if the victim died for causes other than natural, or when the case jurisdiction is outside Bogotá or of a different judicial branch (e.g. military justice). When the forensic pathologist report is required, the case is transferred to the 112 DA of the Life Unit who waits for the pathologist report and, if it is indeed a murder, returns the case to the initial DA for further investigation. Otherwise, the 112 DA closes the case since no crime was committed. Even though there is a specialized DA for femicide cases operating under the previous scheme, the treatment group also investigates these cases if they occurred during a treatment squad shift. We plan to conduct all analysis described below including and excluding femicide cases since these may behave differently.

Third, the DA can meet with his squad to discuss the research strategy and the criminal hypotheses in order to write the methodological program.

Fourth, after the arraignment the case is transferred to a new DA of the Life Unit, who continues with the criminal procedure. This is done so that the DA who investigates the case focuses only on this task and does not spend too much time in court hearings. The homicide squad thus works on the case up to the moment when the bill of indictment is filed. Afterwards a new DA is assigned in order to bring the case to trial. However, we analyze variables beyond this point to see if the changes in the initial stages have effects on subsequent outcomes.

The intervention was piloted for two weeks in December 2015. The intervention officially started on January 20, 2016, and the new investigative procedure was applied to approximately 50% of all new cases. The intervention ended on December 4, 2016.

2.2 Outcome variables

2.2.1 Administrative data

Our main outcomes are actions and decisions taken by the CSIs, detectives, the DA or the judge. Those from detectives and CSIs usually involve activities at the crime scene and additional investigative procedures to find and analyze physical evidence (for example, photographing the crime scene, collecting fingerprints or biological fluids, and interviewing witnesses). The DA's actions are of three types. First, instructing detectives to perform additional investigative actions.

Second, making decisions -like filing charges against a person or arraigning him or her- that define the stage the case is in. Third, attending hearings and presenting petitions to the judge, including orders to interview specific witnesses or asking to deny bail to a suspect. The judges respond to DA petitions and decide whether their actions are legal or not, and assess the evidence and legal arguments presented during the trial to decide on the culpability of the accused.

Some actions are conditional on preceding ones and can only occur at certain stages of the process. Therefore, we will estimate the effect of the intervention on actions at each stage of the homicide investigation process. The set of possible actions or outcomes in a case is extensive and it depends on the characteristics of the homicide. All actions of cases prosecuted by the Office of the Attorney General must be reported by the detectives and the DA in the SPOA (acronym for *Sistema Penal Oral Acusatorio*, also the name of Colombia's criminal justice system). This is a reporting and follow-up automated system, yet officials do not always promptly comply with it. Also, when reporting in the system, they can either choose an action from a menu or write it in their own words, so the same activity might be reported in different ways. Another source of information is the Integrated System of Management of the Judicial Police-SIG (*Sistema integrado de gestión de la policía judicial)*, which is used specifically to report the orders issued by the DA to the detectives.

In addition to the actions and orders reported in the SPOA and the SIG, we will review the initial reports of the crime scene documentation to examine if there are changes in how the investigators describe the scene and in the number and types of activities they perform. Using text mining analysis, we will identify and count keywords and activities that can signal a better investigative process.

There are three challenges in correctly measuring each outcome and estimating the effects of the intervention. First, having a large number of outcomes increases the probability of falsely rejecting the null hypothesis (Romano & Wolf, 2005; Anderson, 2008). Second, data from the SPOA has measurement error since not every action is reported. Third, the same action may be in the system under different labels chosen by officials reporting them.

To deal with these problems, we start by aggregating into a single action all reports of similar activities. For example, actions like finding documents, analyzing databases, study of documents, and all similar actions are classified under a single label called "search and analysis of documents and databases". During this process we drop actions that do not constitute potential meaningful changes in the quality of the process, for instance merely administrative tasks not likely to be changed by the treatment or to significantly change the course of the investigation. We also drop actions for which we lack a clear hypothesis on whether they should be affected by the intervention, for example, indicators of the occurrence of certain control hearings.

We then use the reclassified actions to create summary indices that combine actions to both reduce the number of hypotheses and to produce more precise measures of performance. To create the indices, if necessary we first switch the signs of variables so that increases indicate a better outcome. We convert all outcomes to z-scores by subtracting the control group's mean and dividing by the control group's standard deviation (SD). Finally, we construct the indices as the unweighted average of z-scores for similar actions (Kling, Liebman, and Katz, 2007). Effects can thus be interpreted as mean effects sizes relative to the standard deviation of the control group.

We create the indices only for the first stage of the investigation process because this is where the largest number of actions are undertaken. For the next two stages we focus on the effect of the intervention on individual outcomes.

Because not all additional forensic and investigative actions are properly reported, we focus on those more commonly reported in the data. To define frequency, in the case of index 2 from the administrative data defined in the appendix, we plan to use the distribution of the total number of times each action is reported and drop those that are in lowest decile. We plan to do this without comparing the distribution between treatment and control groups and before doing any estimation of the treatment effect. Doing this limits us to work with the part of the data with less measurement error, which reduces the bias but does not solve the problem completely. Thus, we also plan to compare the actions reported in the SPOA to those reported in the SIG to detect inconsistencies.

After reclassifying actions and creating summary indices we still have multiple outcomes. Thus, we adjust the p-values to account for multiple inference. We follow Romano and Wolf (2016, 2005) resampling-based stepdown multiple testing method to control the family wise error rate (FWER)- the probability of type I error.

Table 1 of the appendix describes stage by stage the list of the outcome variables that will be studied. Outcomes related to the actions performed at the crime scene and to the indictment are of greatest interest to this experiment, thus we will also study them individually, verifying their non-adjusted p-values.

Table 1 presents the descriptive statistics for the outcomes reported on the SPOA for the *pre-treatment* period spanning from December 7 to 20, 2014 and January 20 to October 31, 2015. We used these time periods because they are the same dates in which the intervention takes place in 2016. The SPOA has information for all violent crimes, so in addition to the time restriction we applied the following filters to select our baseline sample: cases in which the initial investigation process was done by the CTI, the crime scene was jurisdiction of Bogotá and the crime was homicide (includes genocide), femicide or abortion.

Table 1. Descriptive statistics
Pre-treatment data

Variables	Mean	Standard deviation	Min	Max	N
First stage: Enquiry					
Number of cases to be established sent to the 112 DA dummy=1 *(Casos por establecer)*	0,492	0,500	0	1	821
Number of cases to be established returned by the 112 DA dummy=1 *(Casos por establecer devueltos)*	0,064	0,246	0	1	108
Unsolved cases *(Archivo por imposibilidad de establecer sujeto activo o pasivo)*	0,118	0,322	0	1	197
Second stage: Prosecution					
Indictment *(Imputación)*	0,070	0,264	0	3	114

	Mean	SD	Min	Max	N
Percentage of cases with indictment (*Tasa de imputación*)	0,109	0,312	0	1	114
Days to indictment	51,298	112,445	0	552	114
Third stage: Trial					
Bill of indictment (*Escrito de acusación*)	0,058	0,240	0	2	96
Percentage of cases with bill of indictment (*Tasa de escrito de acusación*)	0,842	0,366	0	1	96
Days to bill of indictment	114,162	114,681	20	562	74
Trial (*Juicio*)	0,054	0,343	0	7	65
Conviction (*Sentencia condenatoria*)	0,025	0,156	0	1	42
Percentage of cases with conviction (*Tasa de sentencia condenatoria*)	0,368	0,484	0	1	42

Notes: SPOA data for the pre-treatment period spanning from December 7 to 20, 2014 to January 20 to October 31, 2015. Total number of cases is 1667: 568 in the control group and 1099 in the treatment group.

2.2.2 Survey data

A baseline survey was conducted to the CSI teams, the detectives and the district attorneys with the purpose of having additional measures of the mechanisms that may explain the direct effect of the intervention. This survey explores their roles, their motivations and job satisfaction, the quality of their work and the importance of teamwork. A follow-up will be done at the end of the intervention. Using the same methodology described in section 2.2.1 we will create four indices to aggregate several outcomes and we will also evaluate individual outcomes which are listed below:

1. Index 1. This index aggregates four questions about *motivation and feedback*, in which the respondent answers a number between one and ten, where ten is the highest score.
   a) How much feedback do you receive from your superior about your work performance?
   b) In general, how satisfied are you with the support you get from the Office of the District Attorney to do your job?
   c) How motivated and satisfied are you with the work you carry out?
   d) How responsible do you feel for the successes and failures achieved in solving a homicide?

2. Index 2. This index aggregates questions about *role, effectiveness and quality*. As the previous index, the possible answers are the same.
   a) Are the duties that you DA/detective/CSI must develop to solve a murder clear?
   b) In the development of a murder investigation, do you feel you can exercise all the tasks that are assigned to you?
   c) How effective do you think is your team on the ultimate goal of the Attorney's General Office to reduce rates of impunity in the city?
   d) How satisfied are you with the quality of the homicide investigations that you and your team investigate?

3. Index 3. This index aggregates three questions about *teamwork*. The respondent is also required to answers on a scale from one to ten, where ten is the highest score.
   a) When investigating a homicide, how aware are you of the daily tasks that other people of your team are doing?

b) To what extent do you feel that your opinions are valued when making decisions to solve a homicide?

c) Do you feel part of a team?

4. Index 4. This index aggregates questions about *coordination and the progress of the investigation*. The respondent agrees or disagrees with the following nine statements on a five-point scale: strongly disagree, disagree, neither agree nor disagree, agree or strongly agree.

1. The coordination of the team of detectives, CSI and DA in investigating a homicide is good
2. The investigative actions taken to solve homicide cases are often extensive and sufficient
3. The evidence presented by detectives as the basis of the facts of a homicide case presented on his Executive Report is usually sufficient
4. Meetings with the team are usually carried out as often as necessary
5. There is a person in the homicide investigation team who is responsible for effectively coordinating the progress of the investigation and improving its probability of success
6. The dynamics of the current work scheme promotes the emergence of new ideas and useful innovations for the documentation of evidence and case resolution
7. It is important that the DA of the Life Unit knows the investigation in detail from the start
8. When a case is in some sense more complex than the majority of cases, it is easy to contact a specialist on homicides to ask for help
9. It is useful that the detectives and CSI participate in the definition of the criminal hypothesis and orders of further investigation activities

5. *Information*. How much do you agree or disagree with the following statement? "detectives and CSI are not sufficiently informed about the progress and results of their investigation". The respondent answers one of the following: strongly disagree, disagree, neither agree nor disagree, agree or strongly agree.

6. *Overall efficiency*. How much do you agree or disagree with the following statement? "there are often efficiency problems (such as loss of information or evidence, duplication of tasks, wasted work) during a murder investigation". The respondent answers one of the following: strongly disagree, disagree, neither agree nor disagree, agree or strongly agree.

7. *Victims*. Percentage of the time the respondent spends on attention to the victims in a typical work week.

2.3 Qualitative analysis

The homicide investigation is a process that involves multiple civil servants that relate to each other in different ways. These interactions are not captured easily on quantitative data and any changes due to the intervention are difficult to detect. For this reason, a qualitative analysis accompanies the quantitative analysis with the aim of complementing the findings of the direct effects of the intervention. This analysis combines observations and interviews during the intervention period, placing emphasis on their work during a homicide investigation and their relationship with other team members.

2.4 Hypotheses

Our analysis will focus on the average treatment effect to quantify the impact of the intervention. We will focus on Intention-to-Treat estimates, as there might be imperfect compliance with the randomized assignment. Table 1 of the appendix presents the hypotheses for each outcome that will be studied. Panel A describes variables from the SPOA, SIG and the text mining analysis, and, panel B variables from the survey data. The first and second columns of the table contain the outcome names and their definition, and the third column, our hypothesis for each outcome.

2.5 Data

We rely on administrative data on homicides reported to the Office of the Attorney General of Bogotá D.C. Our main data source, from which we obtain our main outcomes, is the Accusatory Penal System-SPOA. This database contains information on all crimes learned by the Attorney General, we focus on the section of felonies against persons such as assault, manslaughter, murder, etc. We also have access to the reports written by the CSI teams, the detectives and the district attorneys for each case of the experiment. These are the corpse examination report, the Executive Report and the methodological program.

Specifically, for each case we have:

- Accusatory Oral Penal System (SPOA)
    - Case id
    - Department, municipality and date of the homicide
    - Department, municipality and date of the complaint
    - Type and seriousness of the crime
    - Anonymized identification of the victim
    - List of district attorneys who have handled the case and date of assignment
    - List of detectives who have investigates the case and date of assignment
    - Crime investigation division
    - List of actions with date

- Integrated System of Management of the Judicial Police (SIG)
    - Case id
    - Type of order issued by the DA and the date of assignment
    - Identification of the detective to whom the order was assigned
    - Date in which the order was fulfilled

- Record of technical examination to the corpse
    - Identification of the victim
    - Number of victims
    - Number of wounded
    - A suspect was captured in the act
    - Description of the crime scene and of the activities done during the examination
    - Description of the examination to the body
    - Description of specific activities such as search for fingerprints, search for biological evidence using forensic lights, video and map of the scene.
    - List of tests requested to the forensic pathologist.

- Executive report

- A suspect was captured in the act
- A person witnessed the crime
- Description of the activities done at the crime scene
- List of the evidence collected at the crime scene

- Methodological program
  - Criminal and investigative hypotheses
  - List of evidence and activities
  - Theory of the case

The survey modules are listed below. Data is available for each person involved in the experiment, but because the control group continued to work with the usual system it is likely that the set of functionaries in this group changed during the course of the intervention. This is because the detectives and CSI are continuously being transferred between units within the Office of the Attorney General.

1. Demographics
   1.1. Gender
   1.2. Age
   1.3. Civil status
   1.4. Job title
   1.5. Years of experience within the Attorney's Office
   1.6. Number of children
   1.7. Highest level of education
   1.8. Job related training.
2. Motivation and feedback
3. Role, effectiveness and quality
4. Team work
5. Work load
6. Coordination and development of the investigation

3 Experimental design

3.1 Aims

As noted in the introduction, the aim of this study is to estimate the effect of a new murder investigation policy adopted by the Office of the Attorney General on several outcomes of the homicide criminal procedure. In order to do so, we use a randomized experimental design that estimates a correct causal effect because it provides comparable treatment and control groups. It is also our aim to assess the mechanisms that could be driving the results through a qualitative analysis of the interactions between the civil servants in the investigation unit.

3.2 Units of analysis

Our unit of analysis is the case. Whether it involves a single murder or more, we treat the entire case as one and we focus on its results. Even though the intervention was designed to treat only murder cases, it is actually the investigation process that determines the felony that the defendant

is charged with. This means that all deaths suspected to be due to causes other than natural ones-excluding car accidents- and that are investigated by the CTI are eligible to participate in the experiment. These include suicide and abortion.

3.3 Randomization

In order to randomly allocate homicide cases between treatment and control, the Attorney General's office had to place strict rotation rules to the investigative teams on call at each shift. The first step was to form fixed investigative units (9 lab experts, 5 detectives and a DA) that would cover the metropolitan area of Bogotá. There were 6 groups in total and 4 of them were assigned to treatment status via a simple raffle. Table 2 of the appendix shows the organization of the treated groups. There are two 12-hour shifts per day and one of the 6 groups is assigned to cover all incidences in each shift. The groups' rotation throughout the month implies that every group covers every one of the 14 possible weekly shifts at least once every 6 weeks (2.3 shifts per week). Since the experiment lasted 42 weeks, each shift was covered by every team approximately 7 times over the course of the experiment. Controls cover each shift at twice the rate since there were 2 control groups (for whom there was no control over the integrity of the composition of investigative teams). Our ITT estimates will therefore compare homicide cases that fall into treatment shifts versus control shifts. Table 2 shows the distribution of treatment groups across possible shifts.

Table 2. Number of times each treatment group covered each type of shift between Jan 20 and Dec 4, 2016.

Shift	Group 1	Group 2	Group 3	Group 4	Control
Monday AM	7	8	7	7	16
Monday PM	8	7	8	8	14
Tuesday AM	7	7	8	8	15
Tuesday PM	8	8	7	7	16
Wednesday AM	8	8	7	8	15
Wednesday PM	7	8	8	8	15
Thursday AM	7	7	8	8	16
Thursday PM	8	7	8	7	16
Friday AM	8	8	7	7	16
Friday PM	8	7	7	8	16
Saturday AM	7	8	8	8	15
Saturday PM	8	8	7	8	15
Sunday AM	8	7	8	8	15
Sunday PM	7	8	8	7	16
Total	106	106	106	107	216

We use baseline data to carry out a placebo analysis in order to test if the randomization schedule alone has an effect on some of the outcome variables describe in section 2.2.1. We apply the randomization schedule described above to the data of 2015 in order to get the distribution of the sample between treatment and control groups. From a total of 1667 cases, 568 would have been randomly assigned to the control status and 1099 to treatment. Following the process described in

section 2.2.1 we create the summary indices in order to test the effect of the randomization mechanism itself. Table 1 of the appendix presents the components of the resulting indices.

Next we regress each outcome variable on the treatment indicator. Table 3 presents the results of the Ordinary Least Squares (OLS) estimation and the raw- and adjusted p-values. After adjusting the p-values we find no significant difference due to the randomization mechanism. Moreover, recall that our indices are standardized and magnitudes indicate that these placebo treatment effects are rarely bigger than 10% of a standard deviation in absolute value an often smaller. Unfortunately, one exception is the percentage of cases with conviction, where the size is large and imprecise and we have a small sample.

Table 3. OLS estimates of the randomization mechanism by stage of the criminal procedure
Pre-treatment data

	N	Treat	Adjusted p-value	Treat	Adjusted p-value
First stage: Enquiry					
Index 1 minimum actions	1.667	-0.00192 (0.0514)	0,999	-0.00353 (0.0421)	0,998
Index 2 forensic and investigative actions	1.667	0.126** (0.0590)	0,348	0.120** (0.0585)	0,585
Number of cases to be established sent to the 112 DA dummy=1 *(casos por establecer)*	1.667	0.0307 (0.0517)	0,993	0.0295 (0.0453)	0,989
Number of cases to be established returned by the 112 DA dummy=1 *(casos por establecer devueltos)*	1.667	0.0555 (0.0535)	0,951	0.0396 (0.0522)	0,989
Unsolved cases *(Archivo por imposibilidad de establecer sujeto activo o pasivo)*	1.667	-0.0545 (0.0506)	0,951	-0.0589 (0.0497)	0,968
Second stage: Prosecution					
Indictment *(Imputación)*	1.667	-0.0109 (0.0504)	0,997	-0.0201 (0.0494)	0,995
Percentage of cases with indictment *(Tasa de imputación)*	1043	0.00623 (0.0652)	0,999	-0.00882 (0.0640)	0,998
Days to indictment	114	-0.0895 (0.190)	0,996	-0.157 (0.185)	0,989
Third stage: Trial					
Bill of indictment *(Escrito de acusación)*	1.667	0.00433 (0.0515)	0,999	-0.00605 (0.0502)	0,998
Percentage of cases with bill of indictment *(Tasa de escrito de acusación)*	114	0.0844 (0.191)	0,996	0.0553 (0.195)	0,995
Days to bill of indictment	74	0.113 (0.245)	0,996	-0.251 (0.259)	0,989

Trial *(Juicio)*	1.667	-0.0153 (0.0512)	0,997	-0.0195 (0.0495)	0,995
Conviction *(Sentencia condenatoria)*	1.667	-0.0562 (0.0489)	0,940	-0.0642 (0.0480)	0,938
Percentage of cases with conviction *(Tasa de sentencia condenatoria)*	114	-0.280 (0.192)	0,803	-0.342* (0.199)	0,778
Controls		No		Yes	
Month FE		No		Yes	
Weekend FE		No		Yes	

Notes: each row presents an OLS regression, robust standard errors in parentheses. Adjusted p-value calculated following Romano & Wolf (2016, 2005). SPOA data for the pre-treatment period spanning from December 7 to 20, 2014 and January 20 to October 31, 2015. The treatment and control groups for this placebo analysis were created by running the randomization schedule of the experiment backwards in time. Total number of cases is 1667: 568 in the control group and 1099 in the treatment group.

4 Analysis

4.1 Regressions

The key parameters to estimate are the direct treatment effects of the intervention on actions taken during each investigation stage. Consider a case c for which a specific procedure y was done, we will estimate the effects via an Ordinary Least Square (OLS) regression:

$$y_c = \beta_0 + \beta_1 Treat_c + \beta_3 X_c + \beta_4 Hours_{lag} + \delta_{month} + \gamma_{weekend} + \varepsilon_c \quad (1)$$

where $Treat_c$ is the experimental condition of the case c and X_c is a vector of case controls. $Hours_{lag}$ is the difference in hours between the date of the complaint and date of death. δ_{month} and $\gamma_{weekend}$ are month and weekend fix effects. ε_c is the robust standard error. Bearing in mind that there are different combinations of lab experts, detectives and DAs in each investigative unit that can be assigned to investigate a case during the urgent acts, and that within these teams exists unobservables that can be correlated, for robustness, we also estimate versions of this model where we assume the standard errors are clustered both at the initial team level[2] and at the shift level.

Because actions depend on court terms, estimates could be contaminated by this differences in time, in order to focus on the differences that come as a result of the treatment as a whole, we estimate the treatment effect on a subset of actions after T number of days have passed since the case was received by the District Attorney's Office. We will define T equal to the 75th percentile of the number of days that pass before action Y is observed, and verify robustness for other thresholds.

4.2 Covariates

In order to increase the precision of the experiment, we plan to include a number of controls such as the number of victims per case and five dummy variables for whether the victim is unidentified,

---

[2] The initial team stays constant for the treatment group, while it may dissolve and change in the controls. However, there might be actions taken initially that influence the overall path of the investigation. This potential correlation within initial teams motivates this robustness check.

if the case is a femicide, if the case is an abortion, if the case was transferred (to another city, to another division within the Office of the Attorney General or to another institution), or if the case started as an assault. Table 4 presents the mean difference between treatment and control groups at baseline for these variables. Additionally, we also plan to include a dummy variable for whether a suspect was arrested *in flagrante*.

Table 4. Mean difference of the control variables
Pre-treatment data

Variables	Control group					Treatment group					Mean Diff
	Mean	SD	Min	Max	N	Mean	SD	Min	Max	N	
Number of victims per case	1,082	0,311	1	3	568	1,086	0,375	1	6	1.099	-0,003
Dummy=1 if victim is unidentified	0,054	0,227	0	1	31	0,058	0,234	0	1	64	-0,003
Dummy=1 if femicide	0,000	0,000	0	0	0	0,001	0,042	0	1	2	-0,001
Dummy=1 if abortion	0,012	0,110	0	1	7	0,012	0,112	0	1	14	0,000
Dummy=1 if the case was transferred	0,044	0,205	0	1	25	0,039	0,193	0	1	43	0,004
Dummy=1 if weekend	0,364	0,481	0	1	207	0,362	0,480	0	1	398	0,002
Dummy=1 if the case started as an assault	0,169	0,375	0	1	96	0,176	0,381	0	1	194	-0,007
Lag in hours between complaint and death	63,48	388,3	0	5.986	568	79,61	408,5	0	5.783	1.099	-16,13

Notes: SPOA data for the pre-treatment period spanning from December 7 to 20, 2014 and January 20 to October 31, 2015. The treatment and control groups for this placebo analysis were created by running the randomization schedule of the experiment backwards in time. Total number of cases is 1667: 568 in the control group and 1099 in the treatment group.

4.3 Power calculations

To estimate the minimum detectable effects (MDEs) of our design, we ran 1000 simulations of random assignment and estimated equation (1). From the total sample of 1667 cases we created a random number and assigned homicide cases to treatment status if that number was equal to or greater than 0.66, the probability of being assigned to the treatment. The original randomization scheme assigned 568 cases to the control group and 1099 to the treatment group. Including the two pilot weeks of December 2015, the intervention lasted 334 days, a total of 668 shifts. For each simulation, we took the standard error of its impact estimate and transformed them into a MDE in standard deviations. We follow Bloom (1995) and multiplied each standard error by 2.80 to obtain the MDEs of a two-sided hypothesis test with a statistical power of 0.80 and a statistical significance of 0.05. We obtained an average MDE of 0.13 SD for the indictment outcomes and an average MDE of 0.14 SD for the bill of indictment, the two main outcomes of the program.

Appendix

Table 1. Outcome variables: definition and hypotheses

	Outcome	Definition	Hypothesis
Panel A: SPOA, SIG, and reports			
First stage			
1.	Index 1: minimum actions.	Set of actions that are usually done in all cases. This set comprises the following actions: search and retrieval of material evidence, interviews, photography of the crime scene, examination to the corpse, documentation of the crime scene, documentation of a location other to the crime scene and verification of things.	We expect no statistically significant difference because minimum actions should be followed in all cases.
2.	Index 2: additional forensic and investigative actions. (Non-adjusted p-value)	Set of additional investigative actions that can be performed in a case. Variables included will depend on threshold frequency as noted in the main text. For the pre-treatment data, the set comprises the following actions: ballistic analysis, search and analysis of documents, physical and chemical analysis, topography of the crime scene, documentation of fingerprints, digital storage and computer analysis, identification of persons of interest through photos and other investigative actions.	We expect an improvement. These actions should be sensitive to the detectives and the CSIs accountability to the DA, which we expect to be affected from the beginning of the investigation in the treatment status, as well as on their cooperation and coordination.
3.	Type of orders issued by the DA to the detectives *(órdenes a policía judicial).*	Orders usually instruct detectives to perform additional investigative actions.	We expect a decrease in orders that are sent recurrently without particular attention to case details, and an increase in orders especially catered to the case.

	If possible, the average fulfillment time will also be studied.		
4.	Number of cases to be established sent to, and returned by, the specialized 112 DA *(casos por establecer)*.	Cases requiring confirmation from the forensic pathologist´s report. If the report confirms the murder, the case is returned to its original DA. Otherwise, the homicide case is closed.	We expect a decrease. Detectives will pay more attention to the crime scene details and to the evidence, thus determining more precisely than the control group if it is a murder case or not.
5.	Number of unsolved cases *(archivo por imposibilidad de establecer sujeto activo o pasivo)*.	Cases without active leads allowing detectives to identify the perpetrator or the victim.	We expect a decrease. A better investigative work will lead to the identification of the suspect and the victim.
6.	Record of technical examination to the corpse *(inspección técnica a cadaver)*		
6.1.	Length of the crime scene description and of the tasks done to document it.	Captures quality of the report: a detailed description gives the district attorney a clearer picture of the crime scene.	We expect more details in the treatment group's description, as this should respond to the quality of the activities done at the crime scene.
6.2.	Keywords count	First respondent: proxy for communication, indicating an interaction with this person about the circumstances in which he or she found the crime scene. Interview: proxy for communication, indicating that the CSI are aware of the interviews done by the detectives. District attorney or detectives: proxy for coordination, indicating that the CSI take into account that their work is an input for the other team members. Photography, video, DNA, ballistics, topography, and number of evidence items	We expect an increase in the frequency of these key words relative to the control group.

		collected: proxy for quality of the crime scene documentation.	
6.3.	Count of additional activities at the crime scene: search for fingerprints, search for biological fluids using forensic lights, videotape or forensic mapping.	Proxy for quality of the documentation of the crime scene.	Depends on the quality of the activities at the crime scene. We expect an increase with treatment.
6.4.	Number of elements sent to the crime lab as evidence.	Proxy for the quality and detail of the crime scene examination.	Depends on the quality of the activities at the crime scene. We expect an increase with treatment.
7.	Executive report *(informe ejecutivo)*		
7.1.	Total number of activities done at the crime scene, and especially number of interviews (which will thus be counted separately as well)	Proxy for quality of detective's work.	Depends on the quality of the activities at the crime scene. We expect an increase with treatment.
7.2.	Length of the description of the tasks at the crime scene.	Proxy for quality of the report	Depends on the quality of the activities at the crime scene. We expect an increase with treatment.
8.	Methodological program *(programa metodológico)*		
8.1.	Number of criminal and investigative hypotheses		The quality of the investigation may lead to a more careful consideration of various hypotheses, but may also help to easily rule out unlikely hypotheses at the outset. We therefore have no clear prediction of the sign of a potential effect, if any.
8.2.	The criminal and the investigative hypotheses are identical	Proxy for accountability	If both hypotheses reported are exactly the same, this might mean that the fields are being filled only to meet a requirement. We expect the treatment squads to investigate the cases in greater detail and therefore have more accurate hypotheses.

18

8.3.	Theory of the case		We expect that a theory is reported and includes keywords such as suspect and witness.

Second stage

9.	Number of cases with indictment *(imputación)* (Non-adjusted p-value)	The suspect(s) was(were) formally charged with the commission of the crime.	Two scenarios are possible: It will increase because of a better initial investigation process and because the intervention eliminates unnecessary transfers between units, which delays the process. It will decrease because the squad homicide spends more time doing a better job processing the crime scene, following leads, and collecting physical evidence allowing them to identify a suspect and indicting only with more solid grounds to ensure a conviction.
10.	Percentage of cases with indictment *(tasa de imputación)* (Non-adjusted p-value)	Number of cases with indictment divided by the total number of confirmed homicide cases. Certain homicide cases are those that were not closed because the absence of a crime.	Same as before.
11.	Days until indictment (Non-adjusted p-value)	For those cases with indictment, time since the initial complaint to get to this point.	We expect a decrease. Elimination of the URI step and the reassignments between DAs should shorten the time of prosecution.

Third stage

12.	Number of cases with bill of indictment *(escrito de acusación)*	A formal written document accusing the suspect(s) of having committed a crime is filed.	May increase or decrease. As a result of a better initial investigation process the DA will have enough high-quality evidence to prosecute

19

			a suspect, indict and write the bill. However, since this action depends on the indictment, the second scenario explained above may also arise, with a more careful job leading to a reduction of potentially spurious indictments.
13.	Percentage of cases with bill of indictment *(tasa de escrito de acusación)*	Number of cases with bill of indictment divided by the number of cases with indictment.	Same as before.
14.	Days until bill of indictment	For those cases with bill of indictment, time since the initial complaint to get to this point.	We expect a decrease. Elimination of the URI step and the reassignments between DAs should shorten the time of prosecution. Although indictments may take longer for the reasons above, we expect the net effect to reduce the total number of days until a bill of indictment is obtained.
15.	Number of cases in trial	Cases presented before a judge for trial.	We expect an increase, but the intervention period might be too short for there to be an effect.
16.	Number and percentage of cases with conviction	Number of cases with conviction divided by the number of cases with indictment.	We expect an increase, but the intervention period might be too short for there to be an effect.

Panel B: Survey data

17.	Index 1	Aggregates questions about *motivation and feedback*	We expect an increase.
18.	Index 2	Aggregates questions about *role, effectiveness and quality*	We expect an increase.
19.	Index 3	Aggregates questions about *teamwork*	We expect an increase.
20.	Index 4	Aggregates questions about *coordination and the progress of the investigation*	We expect an increase.

21.	Information	Proportion of respondents that agree detectives and CSI are not sufficiently informed about the progress and results of their investigation.	We expect a decrease.
22.	Overall efficiency	Proportion of respondents that agree: there are often efficiency problems during a murder investigation.	We expect a decrease.
23.	Victims	Time spent attending the victim's family.	We expect an increase.

Table 2. Treatment group

District attorney	CSI team	Groups of detectives	
		Group 1	**Group 2**
District attorney 1	13	Detective 1	Detective 1
	14	Detective 2	Detective 2
		Detective 3	Detective 3
	15	Detective 4	Detective 4
		Detective 5	Detective 5
		Group 1	**Group 2**
District attorney 2	4	Detective 1	Detective 1
	5	Detective 2	Detective 2
		Detective 3	Detective 3
	6	Detective 4	Detective 4
		Detective 5	Detective 5
		Group 1	**Group 2**
District attorney 3	10	Detective 1	Detective 1
	11	Detective 2	Detective 2
		Detective 3	Detective 3
	12	Detective 4	Detective 4
		Detective 5	Detective 5
		Group 1	**Group 2**
District attorney 4	16	Detective 1	Detective 1
	17	Detective 2	Detective 2
		Detective 3	Detective 3
	18	Detective 4	Detective 4
		Detective 5	Detective 5

References

Alderden, M. & Lavery, T. (2007) *Predicting homicide clearances in Chicago. Investigating disparities in predictors across different types of homicide*. Homicide studies, 11 (2), pp. 115-132.

Anderson, M. (2008) *Multiple inference and gender differences in the effects of early intervention: a reevaluation of the Abecedarian, Perry Preschool, and Early Training projects*. Journal of the American Statistical Association, 103 (484), pp. 1481-1495.

Bloom, H. (1995) *Minimum detectable effects. A simple way to report the statistical power of experimental designs*. Evaluation Review, 19 (5), pp. 547-556.

Brookman, F. & Innes, M. (2013) *The problem of success: What is a 'good' homicide investigation?* Policing and society, 23 (3), pp. 293-310.

Carter, D. & Carter, J. (2016) *Effective Police Homicide Investigations: Evidence from Seven Cities with High Clearance Rates*. Homicide Studies, 20 (2), pp. 150-176.

Casey, K., Glennerster, R. & Miguel, E. (2012) *Reshaping institutions: evidence on aid impacts using a preanalysis plan*. The Quarterly Journal of Economics, pp. 1755-1812.

Cronin, J. et al (2007) *Promoting effective homicide investigations*. Police executive research forum, U.S. Department of Justice Office of Community Oriented Policing Services.

Fiscalía General de la Nación (2009) *Manual de procedimientos de fiscalía en el sistema penal acusatorio colombiano*.

Keel, T., Jarvis, J. & Muirhead, Y. (2009) *An exploratory analysis of factors affecting homicide investigations. Examining the dynamics of murder clearance rates*. Homicide studies, 13 (1), pp. 50-68.

Kling, J., Liebman, J. & Katz, L. (2007) *Experimental analysis of neighborhood effects*. Econometrica, 75(1), pp. 83-119.

National Centre for Policing Excellence-NCPE (2006) *Murder investigation manual*

Romano, J. & Wolf, M. (2010) *Balanced control of generalized error rates*. The Annals of Statistics, 38 (1), pp. 598-633.

Romano, J. & Wolf, M. (2016) *Efficient computation of adjusted p-values for resampling-based stepdown multiple testing*. Statistics and probability letters, 113, pp. 38-40.