Can Trust be Built Through Citizen Monitoring of Police Activity?*

Thiemo Fetzer Caterina Soto Pedro CL Souza

August 13, 2020

Contents

*Fetzer and Souza are based in the Department of Economics at the University of Warwick. Soto is based in the Department of Economics at the London School of Economics.

1 Introduction

New technologies can improve oversight, management of police and, through this process, enhance public confidence and gain citizens' trust. Working in Brazil, in partnership with Santa Catarina state Military Police (PMSC), we study the effect of a program that makes use of new technologies to improve the interaction between the police and citizens. More specifically, we focus on the Rede de Vizinhos ("Neighbor network") community policing program, through which citizens exchange real-time information about crime and public safety through dedicated WhatsApp instant messages groups with the participation of a police officer.

Citizens self-organize and sign up to participate in the Rede de Vizinhos program. As such, it is a necessary condition that prospective participants are aware that the program exists in the first place. In this project, we introduced random variation through an information campaign; our target was to seed group creation and allow us to identify local average treatment effects. This has the benefit of allowing us to identify the effect of the program on the entire state and not just a few select localities. We publicized induction meetings through social media – more specifically, Facebook –, which is an ideal medium to stimulate statewide group creation in a cost-effective manner.

The implementation phase of our project kick-started in June 2018. This was nearly twelve months later than originally anticipated as all projects in this metaketa were held up due to a renegotiation of the EGAP and DFID contract. Ex-post, we identified that up to that point, at least two thousand groups had been formed across the state, a two-fold increase from summer 2017 when we had anticipated originally to commence the project. Our partners, PMSC were optimistic nevertheless that more groups could be formed as a result of the project. Our Facebook information campaign, while reaching at least 700 thousand inhabitants of the state – or, roughly 10% of the population –, did not create substantially more participation in treatment versus control neighborhoods, unfortunately. We presume that this may have been due to the program reaching saturation by the time the research project started. Our baseline surveys indicated that 52% of respondents had heard already about the Rede de Vizinhos program with 13.1% actively participat-

ing. Those numbers, in our view, are commensurate with the interpretation that the program may face decreasing returns to scale. We provide alternative views in our discussion section that may rationalize these outcomes.

We finally present the null results in most hypothesis tests regarding compliance, crime incidence, perception of safety, of the police, police abuse and reporting. Our survey instrument – in line with the other studies and the Metaketa IV questionnaire –, is ample and captures various dimensions of citizen interaction and perception of the police. We do not reject the null hypothesis in all cases (with multiple-hypothesis p-value correction) with relatively narrow confidence intervals. In most cases, our confidence bands lie between -.1 and .1 units of standard deviation. We can thus rule out even mild effects of our intervention.

We proceed as follows. Section 2 provides further details about the context and the Rede de Vizinhos program. Section 3 discusses the intervention and the research design, followed by the results in Section 4. We finalize with a discussion of our results in Section 5.

2 Context

The state of Santa Catarina is in the Brazilian south (Figure 1). With about 8 million inhabitants, it is one of the most economically and socially developed states in the country, with the third highest HDI index and fourth highest GDP per capita, and a relatively well educated population[1].

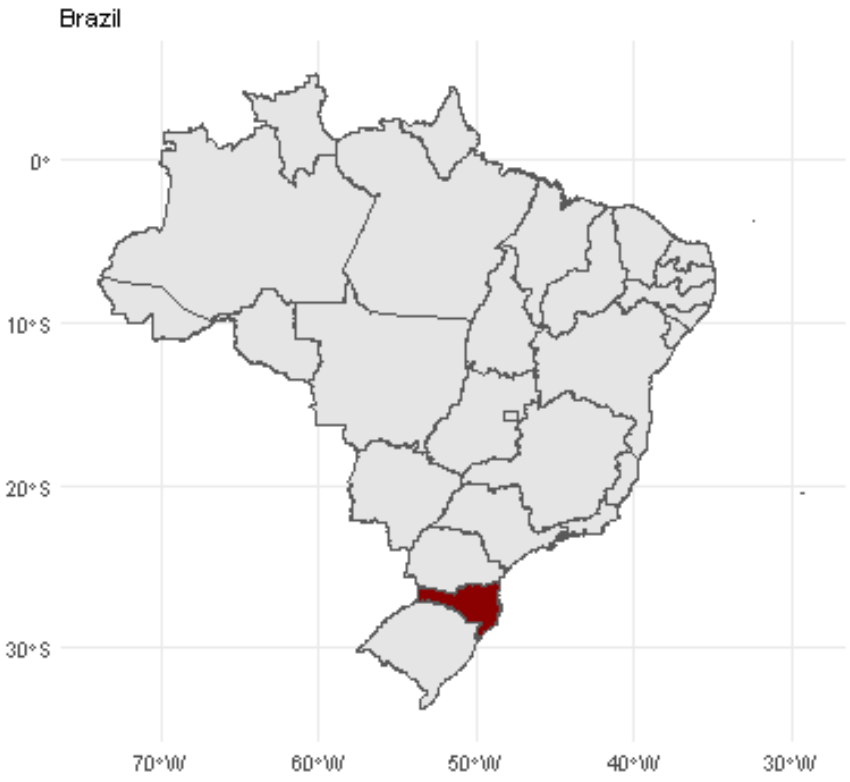
Yet, Santa Catarina struggles with public safety problems and high levels of violence. Brazil is one of the most violent countries in the world, and even states with lower levels of crime have considerably high indicators of violence by international standards. Santa Catarina is the second least violent state in Brazil, and yet the homicide rates stand at 15.2 per 100,000 people in 2017, having increased by approximately 70% up in the previous decade (Cerqueira et al., 2019).[2] Homicide rates are three times higher than the United States and 12 times higher than

[1]Source: IBGE (Brazilian Institute of Geography and Statistics).

[2]Figure 14 in Appendix A shows that there is substantial heterogeneity within Santa Catarina. Most municipalities show relatively low homicide rates, but some show exorbitant levels of violence, specially in the coastal area of the state.

the United Kingdom.

Figure 1: Santa Catarina in Brazil



Since the end of the military government in 1985, a number of community policing programs have been implemented in Brazil to address the policing challenges. The first programs have been carried out by the states of Alagoas, Bahia, Goiás, Mato Grosso, Mato Grosso do Sul, Rio de Janeiro, São Paulo, Rio Grande do Sul and the Federal District (Loche and Mesquita Neto, 2003). The first statewide community program in Brazil took place in the state of São Paulo in 1996. At that occasion, the Military Police of São Paulo and the Royal Canadian Mounted Police initiated an exchange program to discuss experiences regarding community oriented policing (Neto, 2004).

After the military dictatorship ended in the mid 1980s, crime and violence increased considerably across the country. During the dictatorship, the police apparatus in general had a central role in the repression and, up to today, the high

levels of police lethality, brutality and abuse may be seen as an inheritance of the lack of transparency and accountability of that period. During the transition to democracy in Brazil, there were efforts to reform the police towards a more community oriented and humanitarian role. Institutional improvements to increase transparency and accountability were made such as training officers. Similarly, laws obliging police officers to respect citizen rights were enacted (Perez, 1998; Loche and Mesquita Neto, 2003). These changes were important as the population demanded a more transparent and less repressive police at the same time that the crime rates were growing throughout the country.

More recently, Rio de Janeiro adopted another model of community oriented policing program – the so-called Pacifying Police Units (UPP). The program started in 2008 and consisted of a massive increase of police presence – in some cases, the Army and Navy were deployed alongside the regular police patrols –, along with a range of other activities to provide other public goods. This program has been shown to decrease violence in targeted areas (Ferraz et al., 2016; Magaloni et al., 2015). However, the sheer cost and presence of Army and Navy that is required suggests that this model is adequate only for very localized interventions and may not be replicable in a cost-effective fashion across other regions in the country.

Apart from UPP interventions, there has been no rigorous program evaluation of community policing efforts in Brazil: most analyses were limited to comparisons of annual changes in crime indicators based on geographic location, along with qualitative surveys and focus group interviews with program implementation participants.

Rede de Vizinhos in Santa Catarina. The Rede de Vizinhos (RdV) program is a community policing implemented by the Military Police of Santa Catarina in Brazil, established in the first six months of 2016. In the words of the police itself, it aims bring citizens together to *"(...) cooperate with the purpose of fomenting partnerships and strengthening the relation between the police and communities."*[3]

In practice, the program is operated through two layers of contacts between the police and the citizens. First, and perhaps in common with other community

[3] https://www.pm.sc.gov.br/paginas/rede-de-vizinhos, accessed July 14th 2020.

policing programs, in-person meetings are regularly held (often, bi-annually) between the neighbors and a liaison police officer dedicated to the program.[4] The aim of those meetings are primarily to discuss specific security problems of the community and to develop an objective-oriented strategy to address them. They often also involve an exchange of information regarding best prevention practices, recognition and reporting of suspicious criminal activities, and the promotion of a sense of engagement among the local community. Meetings were usually held as close as possible to the group itself – and are often hosted in local venues such as churches, sports hall, businesses or residences with sufficient space to host the group. In fact, the police is flexible regarding the venue selection, as long as it was ample enough to host the group participants. The second layer of communication – perhaps a distinctive feature of Rede de Vizinhos – is the exchange of instant messages via WhatsApp groups which are set up with the purpose of promoting information flows from the police to citizens, from citizens to the police, and among citizens themselves. In addition to these two modes of communication, houses and streets routinely exhibit the Rede de Vizinhos logos as a deterrent device.

Beyond exchange of information, the RdV aims to build a closer relationship between the police and the public, through which trust and cooperation can be built. The component of trust is perceived as particularly salient in Brazil, since cooperation with the police can often carry high private costs such as revenge actions by organized crime groups. Through private communication channels where admissions are highly controlled, the police may be able to increase cooperation with the public. The costs of exchanging information with the police are also reduced by the simple fact that citizens are able to report tips and leads from a mobile phone application. By reducing the costs of reporting – both through reduction of sanctions by criminal groups, and the facility of doing so via a mobile phone application –, it could be expected that the quality and quantity of information reaching the police is far increased, allowing them to allocate their efforts in a

[4]Police officers can be either be fully allocated to the Rede de Vizinhos program – in which case he or she would often allocated to full-time participate in dozens of groups neighborhood – or while in other precincts many officers are assigned one dedicated group and spend only a portion of their time engaging with the group.

more efficient manner.

In parallel, the community policing program aims to improve citizens' perception about intentions and responsiveness of the police. By being able to report back to citizens in a timely manner, the police force would, for example, be able to feed back information on why a particular incident has not been attended to. This could arise for instance if dispatch units are occupied with more relevant incidents. This simple act of reporting back prevents the escalation of citizens' perception about police effectiveness.

The Rede de Vizinhos program could then be expected to address the underlying police-community communication problems through several direct and indirect channels: lowering communication costs (increasing communication from citizens to police); improving the quality of information that is shared; decreasing the chances of revenge acts by gangs; helping to manage expectations about how police will act upon citizen information; opening the opportunity for the police to request information from the public; encouraging citizen-to-citizen communication about suspicious activities in their neighborhoods.[5]

This basic setup highlights that lowering the cost of information sharing may induce citizens to share more information and potentially reduce crime due to its effect on dispatch decisions; however, the setup also highlights a selection margin of the information that gets shared. This could result in information sharing that is noisy and of poor quality, potentially distorted towards crime events that bear the risk of a high private cost (as opposed to public cost).

RdV provides an additional mechanism of communication whereby the police explain citizens their actions as opposed to inactions (which could be a result of the police's information aggregation process or underlying dispatch constraints). Similarly, RdV groups may provide a mechanism for citizens to internalize the (public) cost component of crime, thus increasing the number of information shar-

[5] A theoretical framework of the status quo of police-citizen communications in Santa Catarina would have the following basic ingredients (assuming an exogenous crime generating function): 1) Citizens receive noisy signals about possible crime events; 2) Citizens can choose to report their noisy signal to police at private cost – they will do if the expected (private) cost of a crime event is higher than the cost of sending the signal; 3) Police receives noisy signal(s), aggregate information, and optimally respond (subject to a "dispatch" constraint); 4) Crime and policing outcomes are realized as a function of police activity.

ing events that have a higher externality cost (as opposed to pure private cost) associated.[6]

The process of group creation. The process of creating a new group entails various meticulous verification steps and in-person meetings.[7] As we will clarify later, this process is important to understand the nature and potential effectiveness of our intervention which is based on disseminating information about the existence of the program.

The first step typically involves local citizens – both residents and business owners – coordinating themselves to postulate or apply for a creation of a Rede de Vizinhos group in the place of their residences or commercial establishments. The application is addressed to the local police station. Although the processes substantially vary from place to place, geographical boundaries are usually established at this stage, as well as the initial list of group participants, and group leader is informally elected. Naturally, the step in which citizens self-organize even prior to the existence of the group potentially entails a selection mechanism: it is a necessary condition that prospective participants are aware that the program exists, and that local communities possess some form of social capital, leadership, or organization that enables a joint application in the first place.

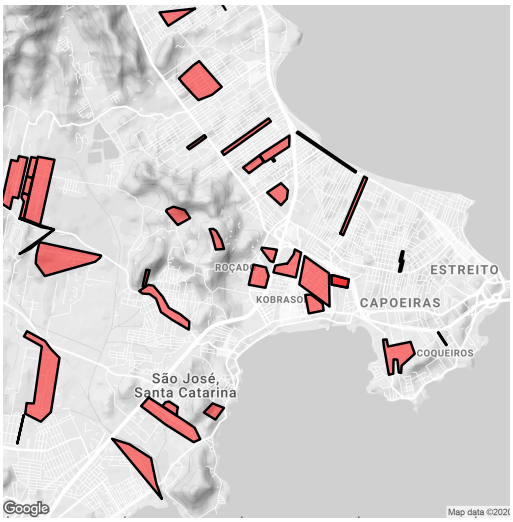
At this stage, the local command has complete autonomy to accept, reject or suggest amends to the proposal. There are no state-wide guidelines regarding the size, number of participants or baseline characteristics of the groups. In the words of police officers in charge of Rede de Vizinhos, groups are "*required to have a sense of geographical intimacy*,"a rather flexible definition. We observed that most likely the definitional flexibility was intended and by design, and in this way preserving the decision-making to local authorities. In practice, the geographic boundaries of RdV groups was left to substantial local discretion. Indeed, Figure 2 plots the spatial coverage of the existing Rede de Vizinhos groups as of June 2018, prior to the start of the intervention, in Florianópolis and Criciúma. It is apparent how the

[6]For more information on the theoretical background, see Section 4 of the meta-study Pre-Analysis Plan, available at http://egap.org/registration/5154.

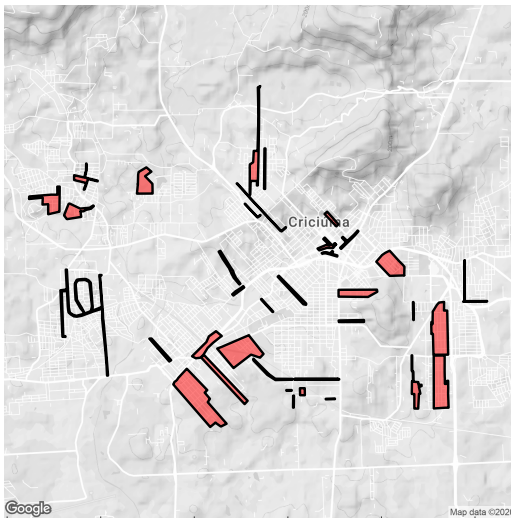
[7]This process is described in https://www.pm.sc.gov.br/paginas/rede-de-vizinhos, accessed Jul 14th 2020

Figure 2: Rede de Vizinhos groups in two selected municipalities (as of June 2018)

Florianópolis



Criciúma



spatial coverage of Rede de Vizinhos could enclose any area from a fraction of a street to multiple streets, or in some cases, entire neighborhoods. Figure 2 also suggests that the exposure to Rede de Vizinhos was quite substantial even before our intervention started.

Individual-level participation is not restricted in any substantial way, provided that individuals are adults with a clean criminal record. As mobile communication via instant messaging groups with a dedicated police contact is a central feature of the program, it is important to note that mobile phone network coverage is reliable across most of the state (and, in particular, across the urban centers). Although there is no formal limit to the number of participants, in practice WhatsApp caps the group participation to 256 individuals. The same number therefore informally applies to Rede de Vizinhos, but this is not a general rule. If a group is accepted, an on-site induction set-up meeting is organized by police officials and local residents. Those meetings are held regularly thereafter as part of the strategy of the program.

Time evolution. As mentioned above, the Rede de Vizinhos program started in 2016 with modest 78 groups which, up to that point, had been created as the in-

dependent initiative of local police precincts and consolidated under the umbrella of the program at its outset. From this point onwards, the project experienced a vertiginous growth. We estimate that the program doubled in size approximately every 200 days or around 6 to 7 months.

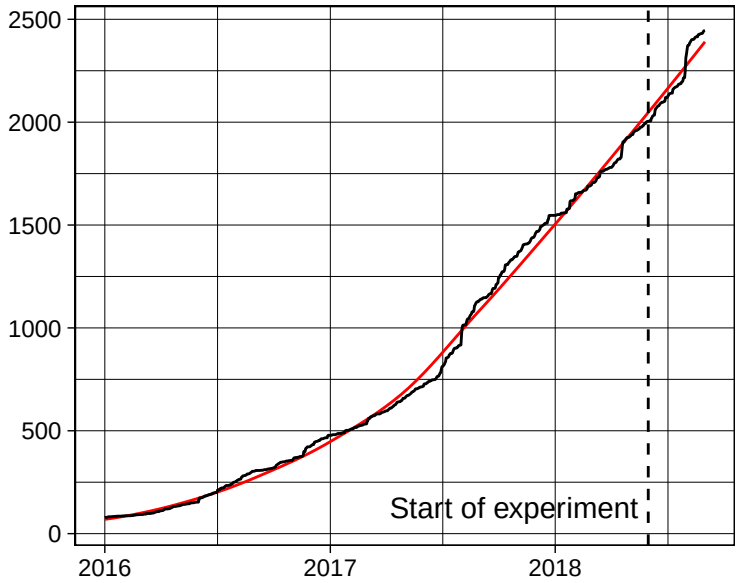
Figure 3 plots an (incomplete) account of the number of established Rede de Vizinhos groups over time. Local commanders were supposed to register the groups under their responsibilities on the "PMSCgeo" system, including the polygon shape of the groups on the map, and basic descriptive information such as the date of creation and number of participants. In practice, not all local commanders registered the groups, so the total number of groups remains unknown. Among those that did, only 81.6% elicited information on when the group was effectively created. We can, nevertheless, use this data to provide insights about the time evolution of Rede de Vizinhos in Santa Catarina.

The initial 78 groups had grown to 478 by the end of 2016, and 1,547 by the end of 2017. Originally, our project was intended to be implemented from late summer 2017. This did not materialize due to delays in the metaketa grant process. By the start of the experimental period – which we here approximate as June 1st 2018 –, over two thousand groups had already been registered. A discrete jump in groups can be observed after the start of the experiment. This can be either attributed to the the experiment itself or to concomitant efforts by PMSC itself to insert all Rede de Vizinhos groups in the system. Taken together, the large number of pre-existing groups as of June 2018 might suggests a saturation of the program (that could also be evident in Figure 2). This is important to interpret the results of the evaluation of the program in Section 4.

3 Intervention and Research Design

As mentioned in the previous section, citizens self-organize and postulate the creation of a group to the police under the original design of the program. This process presumes that *i)* prospective participants are informed regarding the existence of the program; and *ii)* the local communities can achieve an effective coordination to put forth a joint proposal. This could effectively limit the reach of

Figure 3: Number of registered Rede de Vizinhos groups over time



Notes: Number of Rede de Vizinhos groups as registered in "PMSCgeo" management system. Not all existing groups were recorded in the system. Among those recorded, 18.4% did not elicit information on when the group was created. Computed based on the latest data as of October 2018. Red line is the smoothed number of groups. Approximate start of the experiment on June 1st 2018.

Rede de Vizinhos. Instead, we seeded group creation through the provision of an informational campaign delivered through social media. We sought, in this way, to identify local average treatment effects through the variation across space in the localities that would receive the campaign (treatment) against those that would not (control). In practice, we agreed on joint course of action with the police, since not all neighborhoods and localities were suitable to receive the Rede de Vizinhos, or it was not of the police's strategic interest to do so. The purpose of this section is to clarify the design of the intervention and, more specifically, the interpretation of the treatment effects that follows.

This discussion also highlights that treatment effect would be estimated for the subset of groups that are created as a consequence of the information campaign (and would not be created in the absence of the campaign) which effectively might constitute a different selection mechanism from the original design of the program.

3.1 Randomization design and project timeline

We first elicited the interest of local commanders over the entire state in participating on a campaign to create new and increased participation in existing Rede de Vizinhos groups. This early stage of elicitation of interest took the form of an email sent out from the PMSC Headquarters to local commanders across the entire state of Santa Catarina. This occurred in January 2018.

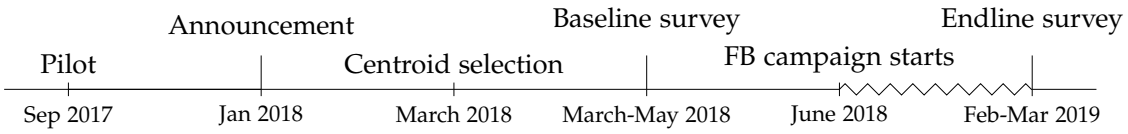
We then contacted the local commanders through phone and WhatsApp to explain the nature of the project, implementation steps and the desire to continue their participation. Local commanders should at all times commit to exchanging information, coordinating actions and keeping the "PMSCgeo" system up-to-date. If positive, we then asked local commanders to select approximately ten centroids. Those points are addresses that could eventually host a Rede de Vizinhos induction meeting. We asked them to select those addresses and send back to us in approximately three weeks' time, up to March 2018. Local commanders would exercise freedom and their judgement to choose those points, and so it is entirely reasonable to expect that focal points to suffer from selection based on unobserved factors. The address roster forms the basis of the randomization.

We then randomly select half of the focal points for treatment taking shape of a Facebook information campaig detailed further below. The other half were allocated to control, stratified at the local command level. This also provides maximum variation at the local level. Given that the randomization occurred after the selection of centroid addresses, it affects both treated and control points equally. That is, the basis of our evaluation is comparing treatment and control focal points that were subject to the same selection criteria.

The randomization was taken at a distance by the research team, immediately after centroid selection. We then later informed the local commanders about the points that were selected for treatment, and asked them to schedule the induction meeting at least six weeks in advance. This delay was necessary to allow for the baseline survey to take place, starting in most cases in June 2018.[8] Endline survey took place in February and March 2019, approximately eight to nine months after

[8]Some meetings could be pushed forward to May 2018

Figure 4: Project timeline
(curvy line represents the implementation period)



implementation started. The timeline of the project is summarized in Figure 4.

Participating police precincts. The request for local commanders to supply candidate points to organize community meetings resulted in a total list of 198 candidate points across 25 municipalities.[9] As explained above, treatment assignment for the encouragement design was such that 50% of the locations within a municipality would receive the treatment. As a result, a total of 100 candidate points were randomly selected for treatment and 98 for control.

Out of the 100 proposed locations that were assigned to receive treatment, 46 dropped out from the study. The vast majority of the locations dropping out (37 of the 46 locations) were lost due to 10 municipalities dropping from the study in full. The remaining 9 lost events were due to various other issues.[10] The full dropping of municipalities means that the treatment and control figures are reconfigured to include 115 locations across 14 municipalities, out of which 58 were assigned to receive the treatment.[11]

The full dropping of municipalities means that our treatment and control figures were reconfigured to include 115 locations across 15 municipalities. Hence the non-compliance to organize actual events that we could advertise was mostly driven by whole municipalities dropping from the study, and not due to selective

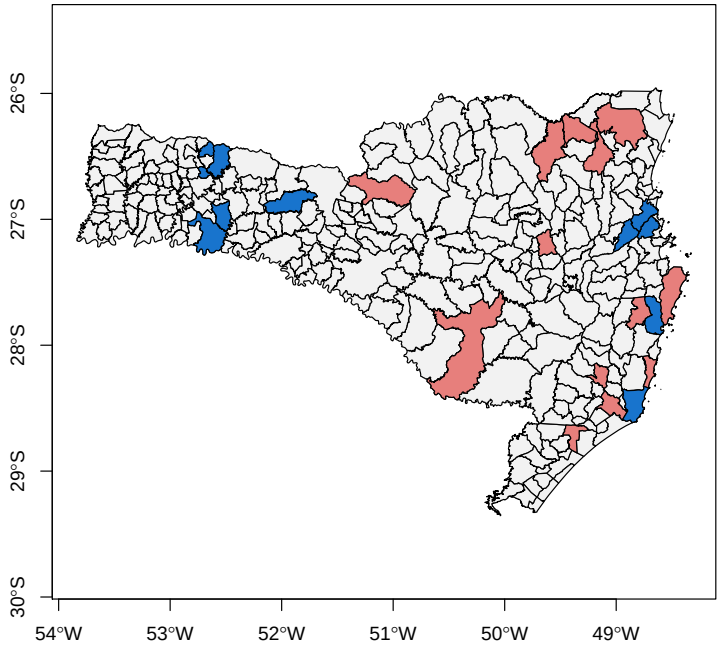
[9]We excluded the municipalities of Coronel Martins, Quilombo and São Lourenço do Oeste because each selected only one point.

[10]Caçador only held 2 events instead of the suggested 5, hence losing 3 T units; Chapecó municipality held meetings prior to the official launch of Facebook campaigns, affecting 5 T units; one event in Tubarão was canceled.

[11]The following municipalities dropped from the study: Brusque (6 points, 3 T and 3 C), Camboriú (3 points, 2 T and 1C), Galvão (4 points, 2T and 2C), Itajaí (10 points, 5T and 5C), Laguna (20 points, 10 T and 10 C), Ponte Serrada (8 points, 4T and 4C), Palhoça (5 points, 2T and 3C), Santiago do Sul (2 points, 1T and 1C), São Domingos (4 points, 2T and 2C), Xaxim (11 points, 6T and 5 C)

Figure 5: Experimental municipalities in Santa Catarina

(delivery = red, no delivery = blue, non-experimental = gray)



non-holding of scheduled meetings. Table 1 shows the list of 25 municipalities that selected candidate centroids, along with the final subsample that delivered *any* Rede de Vizinhos meeting in the "delivery" column. As previously introduced, some municipalities ceased communications and did not follow-up on the project timelines despite positive initial feedback. We provide, in addition, the total number of centroids, and the breakdown by treatment and control allocation. Figure 5 plots the geographical location of the municipalities that initially enrolled in the experiment, that delivered and did not the treatment (respectively, in red and blue).

Table 2 shows that participant municipalities, perhaps as expected, are different to the average municipalities within the state of Santa Catarina. More specifically, column (1) present the statistics for the delivery municipalities and that thus participated in the experiment. For comparison, column (2) refers to the 11 municipalities that registered for participation but did not implement the program. All others are considered non-experimental municipalities in column (3). Finally,

Table 1: Municipalities and Number of Centroids

Municipality	Delivery	Survey	Centroids	T	C
Balneário Camboriu	yes	no	2	1	1
Braço do Norte	yes	no	10	5	5
Caçador	yes	no	10	5	5
Criciúma	yes	yes	10	5	5
Florianópolis	yes	yes	7	3	4
Imbituba	yes	yes	15	8	7
Jaraguá do Sul	yes	yes	10	5	5
Joinville	yes	yes	9	5	4
Lages	yes	yes	10	5	5
Rio do Sul	yes	no	4	2	2
Rio Negrinho	yes	no	3	1	2
Santo Amaro da Imperatriz	yes	no	3	2	1
São Bento do Sul	yes	no	10	5	5
Tubarão	yes	yes	12	6	6
Brusque	no	yes	6	3	3
Camboriú	no	no	3	2	1
Chapecó	no	no	10	5	5
Galvã	no	no	4	2	2
Itajaí	no	no	10	5	5
Laguna	no	yes	20	10	10
Palhoça	no	no	5	2	3
Ponte Serrada	no	no	8	4	4
Santiago do Sul	no	no	2	1	1
São Domingos	no	no	4	2	2
Xaxim	no	no	11	6	5
Total			198	100	98

Notes: Number of centroids submitted by each municipality and for which randomization was drawn. A municipality is considered in the "final sample" if treatment was delivered. "Survey" refers to the whether the centroids were surveyed in the baseline. Note that some municipalities dropped after baseline survey was commissioned. "Centroids" refers to the number of centroids submitted by the municipalities. "T" and "C" refers to the number of centroids randomized into treatment and control, respectively.

we present the p-value of testing the difference of means, comparing delivery vs non-experimental municipalities (and, for the sake of completeness, delivery vs non-experimental municipalities).

We find that delivery municipalities are significantly more populous – on average 140k inhabitants, more than ten times the size of non-experimental places. This is expected as the program is designed, and supposedly more effective, for cities with larger population and more urban. It is indeed confirmed that the share of urban residences is higher in the delivery localities (93.1% vs 57.9%). Those places were also found to be richer, more educated, with more frequent water access, with computer at home and to access the internet more frequently. Interestingly, the characteristics of the non-delivery places are lie between the delivery and non-experimental localities. This once more suggests that self-selection into participation is strongly prevalent. As expanded above, this is expected due to the nature and main features of the program.

3.2 The Facebook experiment

We implement the information campaign through Facebook. The police regularly held meetings for the purpose of disseminating information about Rede de Vizinhos. These meetings could be hosted in town halls, households, or businesses of volunteering citizens, outside of regular working hours. The objective is to hold the meetings as close as possible to citizens' homes to maximize attendance, subject to finding an appropriate location that can host approximately 20 or more citizens. The induction meeting is attended by a police officer who presents Rede de Vizinhos, explains the nature and purpose of the program.

With the collaboration of the police, we publicized the induction meetings through Facebook, only in treatment units. Control group units did not get the social media treatment, and local commanders were discouraged to create new Rede de Vizinhos group in control neighborhoods outside of the experiment.[12] After the definition of the place and time, we posted an "event" on Facebook. Figure 6 shows an example of a social media post made during a successful pilot

[12] In practice, we did not have a way to ensure that groups were not created in control groups. The Facebook informational campaign nevertheless was under the control of the research team and was not introduced in any control point.

Table 2: Comparison between delivery, non-delivery and non-experimental municipalities

	Delivery (1)	Non-delivery (2)	Non-exper. (3)	p-value (1)-(3)	p-value (1)-(2)
Population	140687.7	70439.55	13074.53	0.0010	0.1341
	(148793.7)	(70820.94)	(25083.33)		
Urban (%)	0.931	0.814	0.579	<0.0001	0.0344
	(0.070)	(0.161)	(0.224)		
Income	1563.30	1258.25	1110.38	<0.0001	0.0284
	(416.31)	(224.10)	(222.66)		
White (%)	0.852	0.784	0.830	0.2560	0.0428
	(0.070)	(0.086)	(0.104)		
Primary school or less (%)	0.412	0.519	0.581	<0.0001	0.0026
	(0.075)	(0.081)	(0.075)		
High school or less (%)	0.881	0.925	0.943	<0.0001	0.0069
	(0.050)	(0.020)	(0.021)		
Water access (%)	0.994	0.986	0.987	0.0009	0.0624
	(0.006)	(0.012)	(0.023)		
Computer (%)	0.537	0.428	0.358	<0.0001	0.0135
	(0.098)	(0.105)	(0.107)		
Internet (%)	0.416	0.321	0.241	<0.0001	0.0244
	(0.109)	(0.087)	(0.099)		
Number of Municipalities	14	11	268		

Notes: "Delivery" refers to the 14 municipalities in which the Rede de Vizinhos experiment was implemented. "Non-delivery" refers to the 11 municipalities that registered for participation but did not implement the program. "Non-experimental"refers tot all other municipalities in Santa Catarina state. "p-value (1)-(3)" is the p-value of the comparison of means between column (1) and (3). "p-value (1)-(2)" is the p-value of the comparison of means between column (1) and (2). Income in Brazilian Reais per month. Data from 2010 IBGE Census. Robust standard errors in parenthesis.

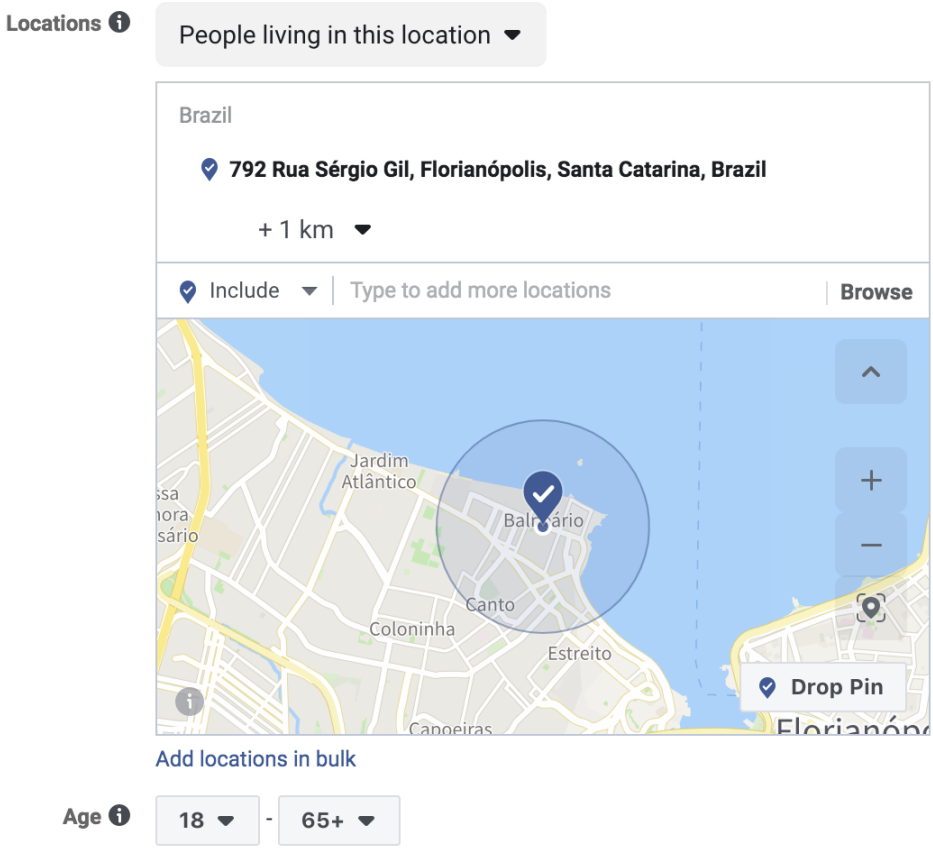
Figure 6: Pilot of the Facebook campaign in Tubarão

communication campaign conducted by the researchers and PMSC in September 2017 in the municipality of Tubarão. The campaign contains the logo of the Rede de Vizinhos program, the date, time and location. Approximately 30 citizens attended that specific meeting. Following the successful pilot, the implementation adhered to this specific format.

To increase the effectiveness of the information campaign (and to help solve the underlying coordination problem among neighborhood residents), we created a dedicated Facebook event for each of our treatment points. The posting of the Facebook materials occurred one week in advance of the actual meeting. We further promoted the event with US$ 10.00 per day, so a maximum spent of US$ 70.00 per meeting.

The information campaign treatment was administered within a tight 1km radius around the "injection points" in which information about Rede de Vizinhos

Figure 7: Targeting of Facebook advertisements



would be disseminated. Figure 7 shows that this can be achieved through Facebook targeting. This was, by design, supposed to contain spillovers, since the advertisement targeting restricted the population that the ads would be shown only to resident population within the circular neighborhood around the centroid (defined by the meeting address). Facebook uses a host of information to infer the likely place of residence, using geo-tagged images and geo-references obtained from mobile usage of the app, resulting in accurate estimates of the proximate home location. We only increased the 1km catchment area if the target population contained less than 10,000 inhabitants. We discuss metrics of engagement with our Facebook campaigns in Section 4.

We now proceed to Section 3.3 and we discuss the data collection efforts and

present the measurement strategy.

3.3 Data and Measurement

Survey planning and data. We apply the survey questionnaire elaborated in collaboration with other teams in the EGAP Metaketa IV funding round.[13] The questionnaire elicited: *i)* demographic characteristics; *ii)* digital literacy and penetration of internet and mobile devices; *iii)* perceptions of the police and cooperation norms; *iv)* victimization; *v)* perceptions about safety and security; *vi)* trust and legitimacy of government, institutions and community; *vii)* exposure to Rede de Vizinhos. The full survey instrument is registered in the Pre-Analysis Plan of the planned meta-study of the EGAP Metaketa IV. To ensure consistent implementation of the survey across treatment and control groups, we aimed to proceed as follows.

First, we selected a set of ten municipalities to receive the baseline survey. The ten municipalities were the most populous among those that had initially registered interest and submitted Rede de Vizinhos centroids, allowing us to conduct the randomization.[14] These are listed in Table 1. Note that two municipalities were surveyed in the baseline and only later dropped from the study (Brusque and Laguna, see Table 1)

Second, for each centroid, we drew 300m radiuses and capped the number of surveyed centroids to ten at the municipal level, randomly selecting five among treated and five among the control. This effectively was only relevant to the municipalities of Imbituba, Tubarão and Laguna. Finally, the target of interviewed households was 68 households per centroid in the treatment group and 34 in the control group. We oversampled from the treatment group to increase the likelihood of capturing households that would eventually participate in Rede de Vizinhos groups. Within the radiuses, the households were uniformly spatially selected. Whenever there were no sufficient number of households in treatment or control

[13] Adaptations of the common Metaketa questionnaire to the Brazil setting, along with justifications are listed in Appendix B.

[14] The sample of municipalities was also constrained by survey logistics which, in practice, disqualified surveying in Chapecó which is located on the west of the state and far from remaining centroids.

areas, all households are interviewed. We then expanded the radiuses to 350m or 400m. If yet there were no sufficient number of households, no further interviews would be conducted at that centroid.

The survey observations were associated to centroids primarily via the GPS reading of the mobile devices where the interview was conducted. In the majority of cases (74.0 and 98.0 percent in the baseline and endline, respectively), this was sufficient to observe precisely where the survey measurement was taken. The remaining 26.0 and 2.0 percent in baseline and endline were geocoded from the provided household addresses. We associated the observation to the closest Rede de Vizinhos centroid among the treatment and control set after filtering for the points that dropped out of the programme.

The baseline survey was implemented between March and May 2018. The window depended on the municipality-specific implementation schedule. Baseline surveying would occur strictly within the approximate seven-week period between centroid selection and the start of the implementation of the Facebook informational campaign, which, in practice, limited the sample sizes, as we expand below. Endline survey occurred approximately eight to nine months after in February or March 2019.

Survey implementation. Figure 8 shows an example of the combined construction of the sampling and centroids in Florianópolis and downtown Criciúma. In those maps, the circles of 300m are drawn in red for treatment centroids, and blue for control centroids. Panel A suggests that some centroids were not completely surveyed and some surveys were conducted outside of the 300m radiuses. Indeed, the surveying was only partially conducted for some centroid due to inclement weather effectively reducing the time window for baseline surveying that was already narrow (within six to seven weeks). This owed to the fact that survey needed to occur between centroid selection and implementation of the Facebook informational campaign. Panel A also shows that some centroids were partially overlapping. As discussed above, we associated the observation to the closest centroid. Panel B shows the example of Criciúma with better sampling properties, ultimately highlighting that surveying efficiency was heterogeneous across munic-

ipalities.

Figure 9 characterizes the sample sizes that were ultimately constructed based on the geographical allocation of treatment. In Panel A, we plot the number of observations in the baseline (dashed) and endline surveys (solid line) against the centroid radiuses. We find 1,514 observations at baseline at the 300m maximum allowable radius and 1,812 observations at the endline. Out of those observations. Panel B plots the share of surveys completed relative to the target (of 68 interviews in treatment, and 34 in control)

All households that constituted the initial baseline sample were to be revisited and prioritised with the purpose of maximising the chances of building a balanced sample. Yet, out of the 1,514 baseline observations, 524 could not be located at the endline, configuring an attrition rate of 34.6 percent. Attrition rates are thus clearly high.[15] We observe that 23.2 percent of attempted interviews were met by an empty household. A further 7.3 percent did not consent to the survey, adding up to 30.5 percent non-response rate. This is consistent with the attrition rates reported above. We do not find evidence of differential attrition comparing treatment and control points.

Administrative data. We obtained the universe of dispatch data from the police for all the municipalities that implemented the Rede de Vizinhos experiment. The data covers the implementation period up to February 2019. The unit of observation is a dispatch, and we have information on: *i)* the time and date of the event; *ii)* the location of the event; *iii)* characteristics of the event and type of crime; *iv)* number of victims, use of force, and arrests. The information is, in most cases, recorded by the police officers themselves through mobile handheld devices. The system automatically embeds the GPS location of the event that triggered the dispatch in the first place and it is therefore accurate in most cases.[16]

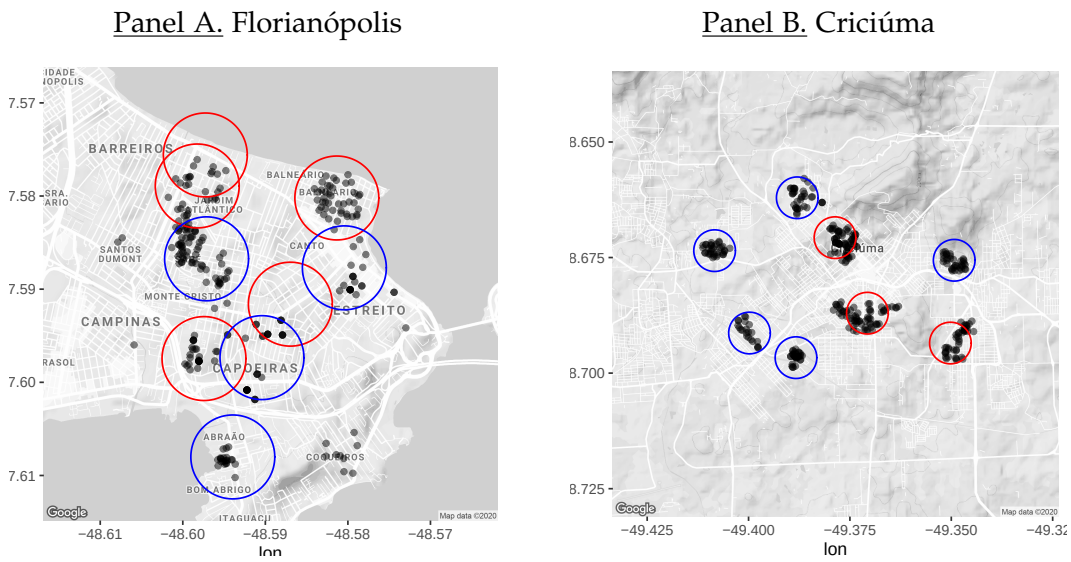
We also obtained information on the shapefiles of the Rede de Vizinhos groups

[15] We established clear protocols to deal with those cases. The protocol stated that household was be visited three times, with no more than two visits in the same day and always at different times of the day. If not successful, the observation would be replaced with the neighbour immediately to the right – or closest to that definition whenever practically possible.

[16] Possible exceptions are the dispatches to investigate possible incidents without clear destination addresses. In such case, PMSC internal systems might register instead a presumed location.

Figure 8: Centroids and Surveying

(red = treated, blue = control)



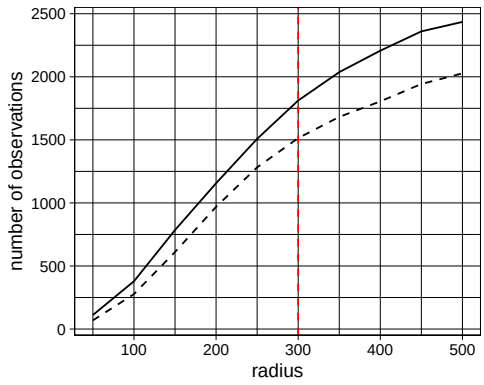
Notes: Survey centroids and 300m radiuses in Florianópolis (Panel A) and Criciúma (Panel B). Red circles are treated, and blue are control. Dots represent a baseline survey observation.

registered in PMSC systems. We obtained two data transfers in June and October 2018. The system was discontinued thereafter and does not provide coverage up to the end of the intervention period. The system does not record the universe of Rede de Vizinhos groups however. While the central command office mandated local commanders to insert data on the system, in reality a share of the existing groups were not inserted and therefore invisible from the point of the police's headquarters. The data therein consists on the polygon of Rede de Vizinhos coverage, the date of group creation and number of participants. We used this data on Figure 3 to construct the time evolution of the number of RdV groups, conditional on reporting into the system.

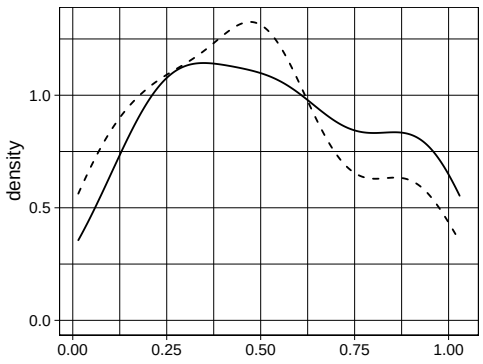
Facebook data. We finally obtain detailed data on the performance of the Facebook adverts. More specifically, we observe the reach (number of unique users that

Figure 9: Survey characteristics

Panel A. Surveys with respect to radius



Panel B. Share of surveys relative to target



Notes: Panel A shows the number of survey observations in baseline (dashed) and endline surveys (solid line) as a function of the radius from the centroid. The red dashed line is the 300m boundary. Panel B is the share of completion of the baseline and endline surveys (respectively, dashed and solid lines) relative to the target of 68 observations for treated points and 34 for control.

visualized the campaign), impressions (number of unique views), social reach and impressions (unique users and views derived from contacts of those that interacted with the adverts), and the number of clicks, likes, engagement, page mentions and views. This allows us to track the effectiveness of the campaign. We drew the data in July and December 2019.

4 Results

Online engagement. The 56 Facebook campaigns achieved a combined reach of 506.0 thousand Facebook users, generating 1.590 million impressions, suggesting that on average, every single user saw our advertisements about 3.1 times. This generated a combined total of 4,768 event responses, or 82.2 per meeting – with recorded attendance in exc706.8ess of 1,075 community members, or 18.5 per meeting –, an effective cost per response close to 93 cents of a dollar per response. Overall, the pages were "liked"4,467 times, and engaged with on 25,185 occasions; individual posts were commented 1,128 times and engaged with 20,718

times (respectively, 20.1 and 370.0 times per meeting). This enabled additional 200.8 thousand individuals seeing the advertisements through the social reach – that is, through the network of friends, or friends of friends engaging with the adverts. Combining the direct and social reach, the numbers indicate that 706.8 thousand individuals in Santa Catarina state saw the ads, or approximately 10.5% of the population of the state. While these numbers suggest that the campaigns were highly efficacious to promote Rede de Vizinhos, we also consider possible that spillover effects may have occurred since our catchment areas most likely are populated by a much inferior share of the population.

Balance and Attrition. We first observe that treatment and control groups were balanced among various characteristics at the baseline. In all specifications, we estimate

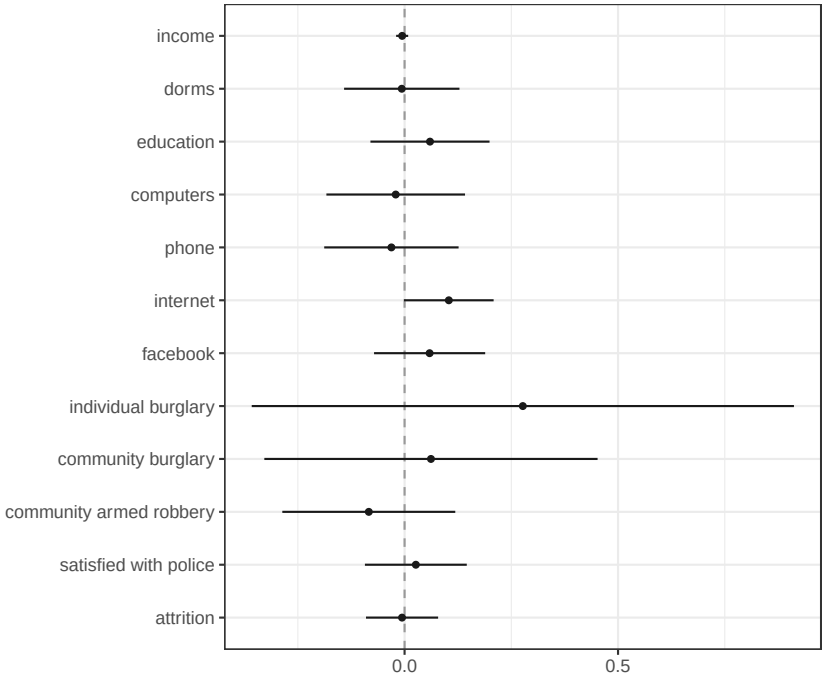
$$y_{icm} = \beta \cdot \text{Treat}_c + \eta_m + \epsilon_{icm} \quad (1)$$

where y_{icm} is the dependent variable of individual i who lives in centroid c in municipality m, Treat_c is the treatment allocation at the centroid level, η_m are municipality fixed effects and ϵ_{icm} is the error term, clustered at the municipality level. We follow the procedures specified in the Meta-PAP and standardize the dependent variable. The coefficient of interest in this case is β.

Our first block of variables consist on demographic features: respondent income in Brazilian Reais, years of education and the number of dormitories, computers, phones in the household. As it can be seen in Figure 10, we observe no significant difference between treatment and control, along any of those dimensions. Our second block of variables are concerned with socio-economic features that could facilitate the access of information, including our Facebook campaign. Again, we find no significant difference between treatment and control relative to internet or Facebook access. This rules out the possibility that treatment effects are driven by the fact that unbalance along those dimensions led to more or less access to the program beyond our intervention. We then test four main outcome variables directly related to security (individual- and community-level burglary)

and perceptions of the police. Once more, we find no evidence that treatment and control centroids were different in their exposure to violence at the baseline. Our final outcome is the indicator whether a surveyed household in the baseline could be located at the endline. As we noted in Section 3.3, this could potentially be an issue due to the relatively high attrition rates that we found in our sample. We note, however, that the coefficients do not indicate any evidence of differential attrition.

Figure 10: Balance and Attrition



Notes: Balance between treatment and control. Coefficients report the coefficient β estimated from Equation (1), along with the 95% confidence interval indicated in the horizontal lines. Labels on the left-hand side indicate the dependent variable. All variables were standardized and coefficients are in units of standard deviations.

Dissemination effect of informational campaign and compliance. We next move to estimating the effect of the information campaign moving markers associated to a wider dissemination of the program. We look at four outcomes for this purpose. We first estimate treatment effects on the frequency in which the police is seen in

the neighorbood, followed by whether respondents reply affirmatively to having heard meetings with police officers in their area. These two questions as indirect and do not allude to Rede de Vizinhos directly. This was meant by design and to prevent any priming effects. Only if positive, we then asked about whether the respondent takes part in Rede de Vizinhos. Only at the endline we asked directly whether the responded knew or had heard about the Rede de Vizinhos. For this latter variable, the mean of the dependent variable is 52.0%, suggesting a wide dissemination of the program beyond our intervention.

We use a similar specification as above, but now include the baseline values of the dependent variable y^0_{icm} for power and an indicator dy^0_{icm} which is equal to one if the observation at the baseline is missing. In such case, y^0_{icm} is attributed a zero value. Again, the disturbance term ϵ_{icm} is clustered at the municipality level.

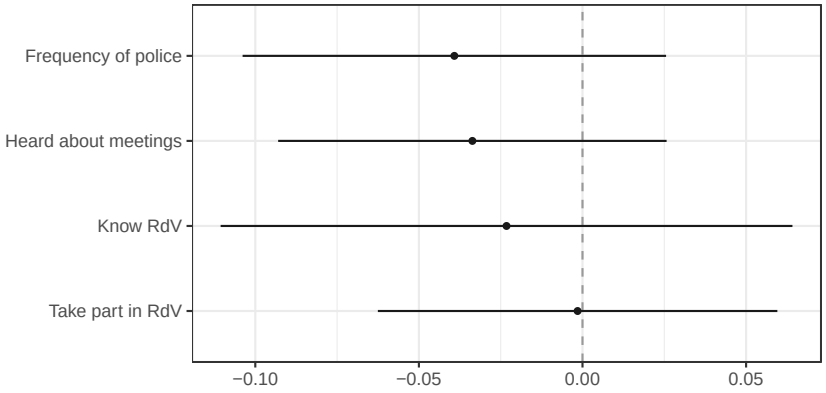
$$y_{icm} = \beta_1 \cdot \text{Treat}_c + \beta_2 \cdot y^0_{icm} + \beta_3 \cdot dy^0_{icm} + \eta_m + \epsilon_{icm} \quad (2)$$

We use this specification for all outcomes, except for "Know Rede de Vizinhos", since this question was only asked at the endline. In this case, we use specification (1) instead.

Figure 11 shows that the information campaign had little effects for all outcomes. A notable null effect is on whether the campaign increased the knowledge or participation in Rede de Vizinhos. We do not observe any effects, with relatively narrow confidence intervals which allows us to rule out effects larger in absolute value to 0.1 units of standard deviation. This result – in light of the dissemination of groups prior to the start of the experiment, and in control groups in the post period –, is suggestive that the program had already reached maturity by the time of our intervention. In this way, an informational campaign had limited scope to further increase compliance or attendance. This may also suggest that the level of engagement that we observed in the Facebook campaign might have been driven by individuals that already had prior exposure to the program.

Outcomes. We next map the treatment effects of the main hypotheses, as described in the Meta-PAP. More specifically, we test for the effects on compliance

Figure 11: Dissemination Effect of the Informational Campaign



Notes: Treatment effects of the informational campaign on markers of knowledge and participation on Rede de Vizinhos. "*Frequency of police*" refers to the question "*How often police is seen in the neighborhood?*". "*Heard about meetings*" refers to the question whether the responded had heard about meetings specifically with police officers in the neirghborhood. "*Know RdV*" asks if respondent knows the Rede de Vizinhos program, and was asked only at the endline. Lastly, "*Take part in RdV*" directly maps individual-level participation in the program. Coefficients report the parameter β estimated from Equation (2), except for the "*Know RdV*"which is estimated by a specification similar to (1). 95% confidence interval indicated in the horizontal lines. Labels on the left-hand side indicate the dependent variable. All variables were standardized and coefficients are in units of standard deviations.

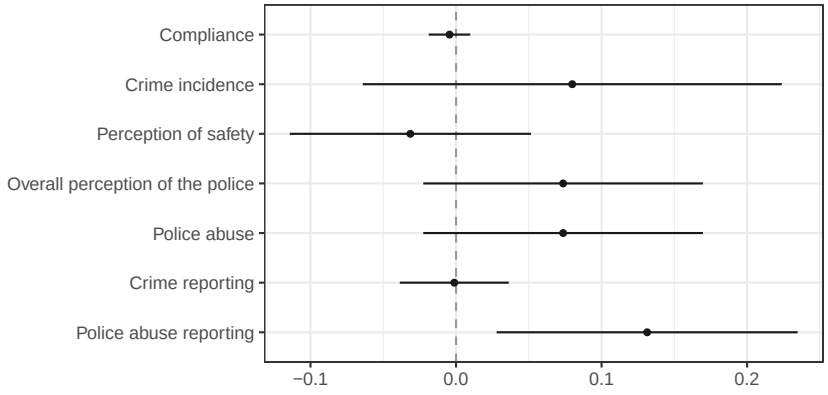
(hypothesis C), crime victimization (1a), perceptions of future insecurity (1b), overall perceptions of the police (2), police abuse (3b), crime reporting (4a) and police abuse reporting (4c).[17] We found no effect in any of the outcomes, with the exception of the police abuse reporting, that turns non-significant with multiple hypothesis adjustment.[18]

We next test the secondary hypotheses and mechanisms: perceived police intentions (M1a), cooperation norms (M1c), perceived police capacity (M2a) and responsiveness (M2b), state legitimacy (S1) and communal trust (S2). We find no effects of the information campaign along any of those dimensions of interest, which is unsurprising given the results concerning the effectiveness of the inter-

[17]We did not implement a test hypotheses regarding police perceptions of citizens (3a) and crime tips (4b) due to lack of data.

[18]The following modifications were made in our study. First, we did not ask for individual-level police abuse due to security concerns and protection of the participants. We also did not ask about the knowledge of the criminal justice system, since the questions in the PAP provided little variation in our pilot sessions and in the early interviews. More details for changes and context-specific adjustments to the Brazil case can be found in Appendix B.

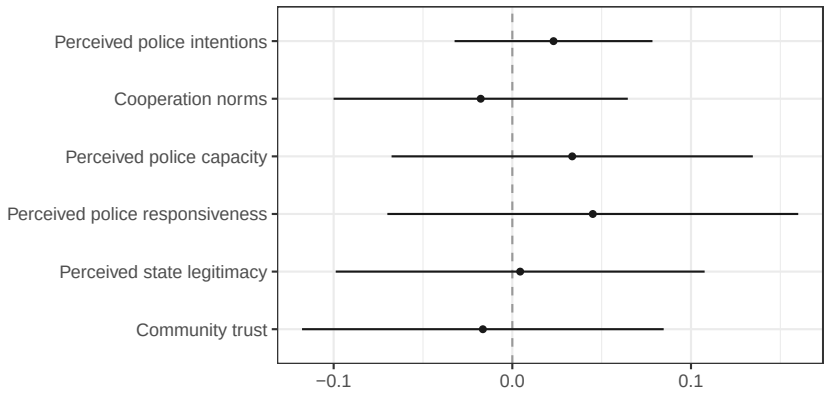
Figure 12: Outcomes: Primary hypotheses



Notes: Treatment effects of the informational campaign on primary hypothesis – i.e., compliance (hypothesis C), crime victimization (1a), perceptions of future insecurity (1b), overall perceptions of the police (2), police abuse (3b), crime reporting (4a) and police abuse reporting (4c). Coefficients report the parameter β estimated from Equation (2). 95% confidence interval indicated in the horizontal lines. Labels on the left-hand side indicate the dependent variable. All variables were standardized and coefficients are in units of standard deviations.

vention and the main hypothesis results.

Figure 13: Outcomes: Secondary hypotheses and Mechanisms



Notes: Treatment effects of the informational campaign on primary hypothesis – i.e., perceived police intentions (M1a), cooperation norms (M1c), perceived police capacity (M2a) and responsiveness (M2b), state legitimacy (S1) and communal trust (S2). Coefficients report the parameter β estimated from Equation (2). 95% confidence interval indicated in the horizontal lines. Labels on the left-hand side indicate the dependent variable. All variables were standardized and coefficients are in units of standard deviations.

We now proceed to conclude and provide a discussion surrounding possible interpretations regarding the treatment effects that we observe in the case of the Rede de Vizinhos program in Santa Catarina.

5 Discussion and Conclusion

The results point to a clear conclusion that the information campaigns did not increase as expected the knowledge or participation in Rede de Vizinhos. This might seem puzzling since, at the same time, the Facebook campaigns proved to be remarkably successful in delivering information in a cost-effective manner. Our findings that over 2,000 groups were active at the outset of the intervention (Figure 3) suggest, in our view, that the program had reached maturity and a certain saturation point (see Figure 2 for the shape files in June 2018 in Florianópolis and Criciúma). The fact that our informational campaigns effectively reached a large swathe of the population can also be driven by existing as well as prospective participants.

We cannot rule out other plausible explanations however. One such explanation involves spillover effects that may have arisen if Facebook targeting effectively failed to contain the information dissemination within the treated and targeted centroids. This may have occurred for a variety of reasons; including endogenous sharing or commenting on the Facebook platform that effectively makes the information salient to others, even if not the primary source of the target. This could also help to explain why knowledge of Rede de Vizinhos was at 52.0% at control points in the endline. As explained above, survey implementation was also subject to delays and some points were under-sampled. We note that the power losses that occurred due to effective smaller sample sizes were effectively of second order since our confidence intervals are small and help to rule out even minor treatment effects.

It is also possible that online engagement does not easily translate into offline action. This might be particularly so since the program effectively presumes a capacity to coordinate actions among local residents in the first place. While we cannot rule out this explanation entirely, we do note that approximately 40% of the

induction meeting participants had known about the event through Facebook.[19]

References

Cerqueira, D., S. Bueno, R. S. de Lima, C. Neme, H. Ferreira, A. Paloma, D. Marques, M. Reis, O. Cypriano, I. Sobral, D. Pacheco, G. Lins, and K. Armstrong (2019). *Atlas da Violência*.

Ferraz, C., J. Monteiro, and Bruno Ottoni (2016). Monopolizing Violence in Ungoverned Spaces : Evidence from the Pacification of Rio's Favelas. *Preliminary Draft*.

Loche, A. and P. Mesquita Neto (2003). Police-Community Partnerships in Brazil. In *Crime and Violence in Latin America: Citizen Security, Democracy, and the State*, pp. 179–204.

Magaloni, B., E. Franco, and Vanessa Melo (2015). Killing in the Slums: an Impact Evaluation of Police Reform in Rio De Janeiro. *Stanford Center for International Development* (556), 1–53.

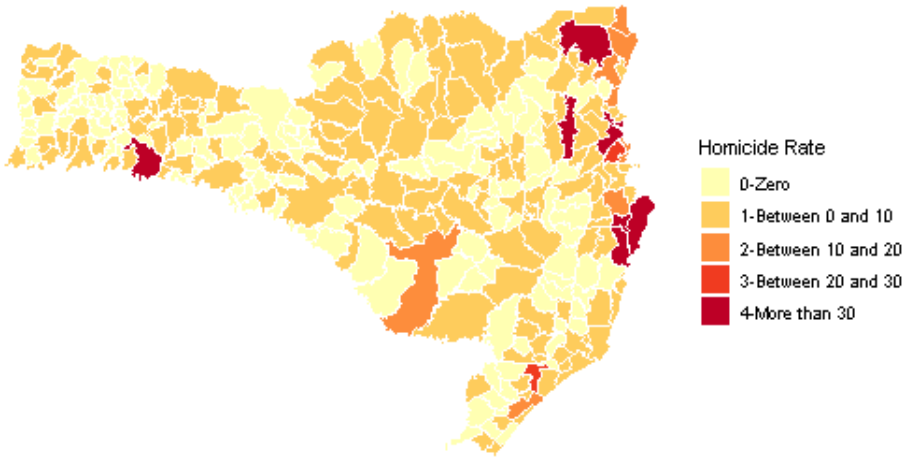
Neto, P. d. M. (2004). La Policía Comunitaria en São Paulo: Problemas de implementacióny consolidación. *Calles Más Seguras. Estudios de policía comunitária en America Latina*.

Perez, R. (1998). 5 Community-Oriented Policing Brazilian Style. In *In War, Those Who Die Are Not Innocent*, pp. 127–168.

[19] This data was collected from incomplete records and meeting attendance sheets.

A Appendix Figures and Tables

Figure 14: Homicide rates by municipality in 2017



B Changes to common questionnaire

The following changes were implemented with respect to the common survey instrument.

1. Simple and aggravated assault: at the baseline we noticed that the reported victimization measures of simple and aggravated assault were very low, and thus this outcome is very likely underpowered (99.1% of the households reported no assault in the six months prior to the survey elicitation). This was partly anticipated and raised due to the ambiguous and unclear translation of an English crime category into Portuguese, as well as the difference in criminal typification across legal systems. Due to space limitations and the fact that the survey was considered to be long, we combined simple and aggravated assault into one measure (assult_num).

2. Domestic violence: after some time in the field, the enumerators stopped collecting data on this question. Enumerators informed the research team about the problematic nature of the question, which resulted in participants ending the interview. The domestic violence questions involve eliciting about a crime possibly in the presence both of the victim and the perpetrator, in the locality where the crime might have happened, which could exacerbate or even provoke further episodes of domestic violence in our setting. As a response, in consideration for our collective responsibility regarding the well-being of the respondents, we stopped asking the question. Yet, respondents were still able to record such episodes through the question on whether he or she was the victim "other crimes".

3. Land disputes and mob violence: Rede de Vizinhos operates exclusively in urban settings where land disputes are not relevant. Also, mob violence is not a relevant feature of our setting. This is further confirmed by the administrative data, where we see zero episodes of either.

4. Police abuse: similar reasoning to simple and aggravated assault - the levels of police abuse were reported to be very low (99.9%) reported no police abuse in their communities. We do not make the difference between verbal and physical abuse.

5. Report norms: reportnorm_theft and reportnorm_abuse replaced by a more general "Interacting with the police may not be well seen by community members" in polint_comm.