Relational State Building in Areas of Limited Statehood: Experimental Evidence on the Attitudes of the Police

SABRINA KARIM *Cornell University*

*U*nder what conditions does state expansion into limited statehood areas improve perceptions of state authority? Although previous work emphasizes identity or institutional sources of state legitimacy, I argue that relationships between state agents and citizens drive positive attitude formation, because these relationships provide information and facilitate social bonds. Moreover, when state agents and citizens share demographic characteristics, perceptional effects may improve. Finally, citizens finding procedural interactions between state agents and citizens unfair may adopt negative views about the state. I test these three propositions by randomizing household visits by male or female police officers in rural Liberia. These visits facilitated relationship building, leading to improved perceptions of police; shared demographic characteristics between police and citizens did not strengthen this effect. Perceptions of unfairness in the randomization led to negative opinions about police. The results imply that relationship building between state agents and citizens is an important part of state building.

INTRODUCTION

In areas of limited statehood, a country's central authority lacks the ability to implement and enforce rules and decisions or lacks legitimate monopoly over the means of violence (Risse and Stollenwerk 2018, 405). State building is the act of expanding the state's central authority into these spaces, so that the state's institutions, including its enforcement arm (police), become the dominant source of goods and services provision (Krasner and Risse 2014; Lake 2016). For the state to develop this monopoly, citizens must confer legitimacy onto the state and its (enforcement) agencies (Risse and Stollenwerk 2018). The state earns

legitimacy when the public accepts state authority over other sources of authority and as this state authority remains perceptually favorable to the public (Braithwaite and Levi 2003; Levi and Sacks 2009; Levi, Sacks, and Tyler 2009; Risse and Stollenwerk 2018; Schatzberg 2001; Tyler 2006). As the state expands its (enforcement) capacity, and as individuals engage with new state agency personnel, citizens' attitudes toward state authority may change. Under what conditions, then, does state expansion into areas of limited statehood improve perceptions of state authority?

Previous work has primarily relied on institutional or identity-based sources of legitimacy (Risse and Stollenwerk 2018, 410), but their application is only possible for certain segments of the population. Instead of (or in addition to) institutional or identity-based explanations, I argue that as the state expands into areas of limited statehood, individual relationships between state agents and citizens are crucial for understanding how legitimacy is formed. As individuals build relationships with state agents, such relationships may help improve perceptions of state authority through two mechanisms. Using "contact theory" as a starting point (Allport 1954), I suggest that relationship-building interactions between a state agent, who is a representative of state expansion, and a citizen: (1) lead to positive updating of negative priors and (2) facilitate social bonds between the agent and the citizen (Pettigrew 1998; Pettigrew et al. 2011; Pettigrew and Tropp 2006). The *relational contact mechanism* thus posits that as individuals learn new information about the state agents and as they develop social bonds with state agents, locals may be more likely to accept the state's authority.

Information and social bonds develop more quickly when the state agent and the citizen share visible demographic characteristics such as the same ethnicity, race, or sex. This is because outwardly visible demographic characteristics make it easy to recognize a member of the organization who may have concern for people who look similarly. The signal of similarity could also more easily facilitate common ground during a social interaction. Indeed, bureaucratic representation theory

Sabrina Karim [iD], Assistant Professor, Cornell University, smk349@cornell.edu.

posits that when state agents look like citizens, policy outcomes improve (Meier 1975; Meier and Nicholson-Crotty 2006). This would mean that when female agents of the state interact with local women, *local women's* views about state authority should improve.

Finally, interactions between state agents and citizens become transactional or procedural as the state performs its duty in new territories (e.g., enforce the law, rule on court decisions) or as it provides goods and services (e.g., cash transfers, social security IDs). If these transactions with the state are perceived as unfair, the experience could undermine legitimacy (Mcloughlin 2018; Sunshine and Tyler 2003; Tyler 2006). For "procedural justice" to induce positive perceptions of the state, citizens must perceive state agents to be neutral, unbiased, and fair in their transactional interactions with citizens (Tyler 2006). The *procedural justice mechanism* thus posits that when citizens' procedural interactions with state agents are perceived as fair/unfair, they will develop positive/negative views about the state's authority (Sunshine and Tyler 2003; Tyler 2006).

To test the relational contact mechanism, I conducted a field experiment in one of the most remote counties in Liberia: Grand Kru County. I partnered with the Liberian National Police (LNP) to randomize face-to-face household visits by police officers as a way to expand their community policing program into an area of limited statehood. In addition to randomizing the household visits, I randomized the sex of the police officers, which allowed me to test whether shared a demographic characteristic (in this case, sex) enhanced perceptions. In total, 375 randomly assigned households received 20–30 minute visits by two male police officers, 375 randomly assigned households received visits by two female police officers, and 225 randomly chosen households did not receive visits.

The results from the experiment show that household visits led citizens to prefer the police to provide security during a crisis over non-state and external entities, and it showed that citizens were less likely to perceive the police as abusive and more likely to perceive them as effective. At the same time, I do not find a heterogeneous treatment effect based on sex. The findings imply that initial relationship-building interactions between the police and citizens improve the perceptions of state authority, at least in the short run. Sharing a demographic characteristic with state agents, however, does not appear to lead to a larger effect.

During the experiment, some members of the control group heard about the household visits. This group believed that they were unfairly deprived of a state-provided service. Consequently, the spillover allows me to observationally test the effects of the procedural justice mechanism. I find that control group members who heard about the treatment that others received developed negative attitudes toward the police. This finding not only raises concerns about valid inference, but it also requires reflection over the ethics of experiments and policy decisions that do not provide services to everyone if it is possible to do so.

This paper makes several important contributions to the literature on the sources of legitimacy, public opinion formation, state building, and gender. Although existing studies have shown the importance of understanding the nuances around performance as a source of legitimacy (Braithwaite and Levi 2003; Schmelzle and Stollenwerk 2018), this study joins others in demonstrating the importance of relationship building as a part of the state-building process (Fisk and Cherney 2017). It shows that the formation of relationships between state agents and citizens plays a key role in legitimizing state institutions during state expansion. Moreover, the study shows that public opinion is malleable and can be shaped based on citizen–state agent relations. Recent work on the state building literature using randomized controlled trials has found that interventions have an effect on behavior, but that it is more difficult to change perceptions (Beath, Christia, and Enikolopov 2013; Paluck 2009). This study shows that, under some conditions, changes in perception and public opinion are possible. As such, the results are consistent with other work related to door-to-door canvassing (Broockman and Kalla 2016).

Methodologically, the paper provides guidance on how to ethically do experimental work in contexts of state expansion. It also provides a new way to experimentally test the effect of interactions with people of different sex on outcomes of interest while holding personality constant. Finally, the paper highlights how state expansion could proceed in the future. State builders should take care to ensure that state agents who enter into areas of limited statehood engage in relationship-building activities with locals and that procedural interactions between state agents and citizens are perceived as fair.

STATE BUILDING IN THE CONTEXT OF LIMITED STATEHOOD

Areas of limited statehood are defined as areas where a country's central authority is to some degree unable to implement and enforce rules and decisions (Risse 2011; Risse and Stollenwerk 2018). The country's central authority does not have legitimate monopoly over the means of violence. The state lacks what Krasner (1999, 4) calls "domestic sovereignty" or the "formal organization of political authority within the state and the ability of public authorities to exercise effective control within the boundaries of their own polity."[1]

Importantly, areas of limited statehood do not always suffer from a deficit of governance, or rather, they do not always lack institutionalized modes of social coordination that enforce rules and provide collective

---

[1] It is similar to what Mann (1984) calls the "infrastructural power of the state." Infrastructural power is the aspect of the state that determines how far its bureaucracy can reach to exert control and regulate social relations (Soifer 2008, 234). The definition is also in line with Weber (1980) for whom statehood is conceptualized as the state's ability to rule authoritatively and to legitimately control the means of violence.

goods and services (Risse and Stollenwerk 2018, 406). Non-state actors such as traditional leaders or even insurgent/criminal networks often provide goods, services, rules, and structure to societies in areas of limited statehood, as do external actors, such as United Nations peacekeepers, aid agencies, and even multi-national corporations (Arjona, Kasfir, and Mampilly 2015; Brass 2016; Risse 2011; Schmelzle and Stollenwerk 2018).

The act of state building expands the state's central authority into these spaces, which could lead to competition for control between the state and non-state or external "governors." Often, the state's first area of expansion is through security provision (Lake 2016). This means that the state starts to build new police stations and sends individual police officers to these territories (Blair, Karim, and Morse 2019). Police officers start to replace indigenous security institutions such as traditional leaders or civilian defense forces. In some post-conflict contexts, they replace peacekeeping missions.

For the state to develop sole authority over these territories, citizens must confer empirical legitimacy onto the state and its agencies (police) (Risse and Stollenwerk 2018). Empirical legitimacy (hereafter legitimacy) means that the population develops an obligation or willingness to accept the state's authority in their territory (Risse and Stollenwerk 2018; Levi, Sacks, and Tyler 2009; Levi and Sacks 2009, 404). This can be conceptualized in a dyadic manner, whereby the public prefers goods and services provided by the state over goods and services provided by non-state or external actors. It can also be conceptualized in a monadic manner, in which citizens evaluate government goods and service delivery without comparison. They may make their evaluation based on whether state agencies violate citizens' sense of what are acceptable and appropriate modes of behavior (Risse and Stollenwerk 2018; Schatzberg 2001). At minimum, a violation of acceptable behavior occurs when state agents engage in abusive, offensive, or corrupt behavior, or if they provide goods and services in an ineffective way. As such, citizens' perceptions of the state agents' behavior are important for understanding their legitimacy. I use these two conceptualizations—a preference for state agencies over non-state actors and external actors and opinions about the behavior of state agencies—to measure legitimacy.

State building, or the act of expanding the state's central authority into areas of limited statehood (Krasner and Risse 2014; Lake 2016), has the potential to change citizens' attitudes toward state authority. In other words, state expansion may, in and of itself, have a legitimizing or delegitimizing effect. Previous work has provided identity-based explanations for the sources of legitimacy, such as charismatic leadership, ethnic belonging, or legal recognition (Risse and Stollenwerk 2018). Although these are important sources of legitimacy, they apply to certain parts of the population, such as those who value charisma in a leader or who are ethnic minorities.

Risse and Stollenwerk (2018, 410) also provide institutional explanations for legitimacy, such as input-based participatory governance or output-based performance (effective goods and service delivery). The former suggests that democracy or fair elections should legitimize the state, yet state building efforts that prioritize elections over other institutions have often failed (Paris 2004; Snyder 2000).[2]

The performance-based mechanism posits that when the state's expansion leads to the effective delivery of goods and services to those in the territory, perceptions of the state may improve (McLoughlin 2015; Zartman 1995; Rotberg 2010; Schmelzle and Stollenwerk 2018; Ziblatt 2008; Weber 1962).[3] Yet, the performance-based legitimacy mechanism requires shared goals between citizens and the state, and it requires attribution of the governance's success or failure to the state agency (Schmelzle and Stollenwerk 2018; Stollenwerk 2018). The public goods and services preferences for those in the territory must align with what the state provides. Even when there is alignment and the state effectively delivers goods and services, citizens may not attribute the change to the state (McLoughlin 2015). Moreover, expanding goods and service provision could raise expectations about the state's ability to provide goods and services, which, if unmet, could lead to resentment (Blair, Karim, and Morse 2019; Nussio et al. 2019).

Finally, Risse and Stollenwerk (2018, 410) suggest a third source of legitimacy, which is based on social trust or esteem of authority. They refer to trust and esteem developing for non-state actors, such as religious leaders, traditional chiefs, neighbors, or local leaders. These actors are likely to have moral authority over the public due to a shared or collective set of values, or due to the high level of face-to-face interactions (Schatzberg 2001, 24). I argue that this form of personalized trust or esteem can develop for state agents and not just local (non-state) actors. As such, I suggest a relational approach to state building in which legitimacy may form through the development of personal relationships with representatives of the central state who are responsible for implementing goods and services to the public.

ATTITUDE FORMATION IN AREAS OF LIMITED STATEHOOD: A RELATIONAL APPROACH TO STATE BUILDING

Before exploring how relationship building helps legitimacy form, I first describe two factors that may influence attitude formation in post-conflict countries, as my measures of legitimacy are attitudinal. First, citizens' prior attitudes toward the police affect whether or not increases in police presence change attitudes toward the police. Messages received early in life affect people's attitudes toward a given topic, and these early attitudes affect how subsequent information about that topic is

---

[2] Nussio et al. (2019) find that participatory governance could have adverse effects on perceptions of the state.
[3] See special issue by Schmelzle and Stollenwerk (2018).

evaluated (Chong and Druckman 2012). Opinions may form due to experiences, information gained from radio broadcasts, newspapers, word of mouth, or rumors. In areas of limited statehood, even though individuals may have few experiences with the police, they may still develop opinions about them because of exposure to information about the police. These opinions, however, are likely to be negative. In general, few people think fondly of the police as their encounters with them usually mean that something bad has happened (Huo and Tyler 2002, 5). Even if people have infrequent or no prior interactions with the police, they may have heard about other people's bad experiences or heard rumors of abusive police behavior.[4] Moreover, those living in post-conflict areas may have experienced state-perpetrated abuse. Such predatory behavior by the state likely generated negative perceptions of the police. Thus, I assume that individuals' priors about the police are generally negative.

Second, individuals' predispositions to form opinions affect how they process new information. Individuals process new information in different ways; they vary in their predisposition to form spontaneous evaluations when presented with new information about a topic. One group of people develop immediate opinions when presented with new information; they are called "need to evaluate" individuals (Chong and Druckman 2010, 666). In contrast, "online or memory-based" individuals require some form of recall (e.g., surveys) about the topic to form an opinion (Chong and Druckman 2010, 666). Although "need to evaluate" individuals tend to develop stronger opinions, the distinction is irrelevant in this study because the survey questions prompt the "online or memory-based" individuals to think about the police and to develop an opinion. As such, in this study, I assume that any individual can form or change their opinions about the police when presented with new information about them.

The Relational Contact Mechanism

Contact, or an extended, face-to-face social interaction, between two antagonistic groups can promote tolerance and acceptance under certain circumstances, such as when groups have equal status, a common goal, cooperation, and authority support (Allport 1954). Since the development of contact theory, it has been generalized to apply to many different types of groups (Pettigrew and Tropp 2006), including the police (Rosenbaum et al. 2005; Viki et al. 2006). Moreover, the theory has held even when Allport's (1954) conditions are not met (Pettigrew and Tropp 2006). I apply particular mechanisms of contact theory developed by Pettigrew (1998) to the context of areas of limited statehood and citizen–state contact.

Pettigrew (1998) develops contact theory by exploring the processes through which contact achieves improved perceptions and behavior toward an out-group or toward a social group with which an individual does not identify. In particular, Pettigrew (1998) stresses that improvements in perception occur when contact takes on a relationship-building form. Relationship building means developing a mutual understanding or connection with the out-group member (police, in this case). The first interaction or contact is important for the development of a connection as it sets the stage for future interactions. Pettigrew (1998) posits that perceptions change to become positive on the first interaction if (1) interactions lead to positive updating of negative priors and (2) if they facilitate social bonds.[5]

First, he suggests that contact with out-group members provides new information to the person that could update his or her prior beliefs about the out-group. Citizens learn new information about the out-group member as they interact with and observe the behavior of the out-group member. Individuals' perceptions change when the out-group member's behavior is starkly inconsistent with their prior negative beliefs about the out-group. As mentioned above, priors about the police (the out-group) tend to be negative. But if the police act in a manner that is inconsistent with people's priors—they act friendly instead of abusive—then citizens may update their beliefs about the police to be more positive. In contrast, if the police continue to be predatory, then no updating occurs or the existing negative perceptions harden (Chong and Druckman 2010).

In addition to information updating, interactions could generate social bonds between in-group and out-group members (Pettigrew 1998). Social bonds form as initial anxiety about the interaction subsides over the course of the visit. Pettigrew (1998) suggests that longer encounters and repeated encounters allow social bonds to form and deepen. Other studies have found that repetition may not be needed and that a single encounter can lead to empathy. Broockman and Kalla (2016), for example, find that during door-to-door canvassing, single ten-minute conversations reduced prejudice of transphobia for at least three months. This research suggests that extended interactions facilitate dialogue and empathy, which could humanize both actors and subsequently reduce anxiety.

Moreover, prior to the visits, opinions about the police may have formed in the abstract, but the door-to-door visits put a literal face to the police. Social bonds, after all, only form with an actual person. After the visit, when people think about the police, they are likely to think about the individual police officer that they met and about the interaction they had with that person. As such, the nature of that (one) interaction—whether it is information updating or leads to social bonds—matters for generating positive opinions about the police.

Yet, even if police interactions facilitate positive perceptions of individual police officers, these positive effects may not extend to the entire organization.

---

[4] Rumors about state security officials are usually always negative. See Greenhill and Oppenheim (2017).

[5] Pettigrew (1998) does not specify whether one or both processes are necessary for contact theory to apply. Similarly, in this paper, I do not test each mechanism individually but rather facilitate a relationship-building interaction that includes these two elements.

However, Pettigrew ([1998](#)) argues that the perception of the individual police officer generalizes to the entire group if the police officer is a "typical representative" of the group. In this case, a "typical" police officer is one who wears a police uniform. The uniform is used to suppress individuality and designate membership in a group (Joseph and Alex [1972](#)). When wearing a uniform, the police officer adopts the uniform identify of the police. Thus, as long as police officers wear uniforms, they may be considered "typical" members of the group.

Importantly, "typical appearances" of police officers differ from the "typical behavior" of police officers. As mentioned above, citizens may understand the typical behavior of police officers to be abusive. As such, atypical, non-predatory behavior of police officers during social interactions could facilitate changes in opinion precisely because it is atypical. The uniform, however, demonstrates that this "new" behavior could be typical of other police officers in the organization.

In short, when individuals living in areas of limited statehood build relationships with police officers, they may be more likely to perceive the police positively because the encounter elicits perceptional change about the individual police officers with whom they met, and this perceptional change also generalizes to the police more broadly. This leads to the following hypothesis:

Hypothesis 1. *When citizens living in areas of limited statehood engage in relationship-building interactions with police officers, they are more likely to develop positive perceptions of the police force.*

Strengthening the Relational Contact Mechanism

Although Hypothesis 1 extends to any police officer, it is possible that certain types of police officers may be more likely to elicit positive perceptions than others. In particular, when police officers share demographic characteristics with the population or, more bluntly, when they look like the population, they may be more easily able to develop relationships with locals. The idea is similar to Pitkin's ([1967](#)) concept of descriptive representation, in which political representatives, by virtue of looking like the population, are accorded some form of legitimacy. Mosher ([1982](#), 12) applies theories of descriptive representation to state agencies, suggesting that a state agency is "passively representative" if the composition of the police force mirrors the population in terms of race, social class, ethnicity, religion, gender, occupation, or other such characteristics.

Given the public's negative priors about the police, the inclusion of police officers who look like citizens into the police force could signal that the police are willing to serve the minority (Meier [1975](#); Pitkin [1967](#)). The visible, shared demographic characteristic of being female makes the female officer easily identifiable as someone who may have concern for other women or members of the minority group. Women may feel that female officers are less likely to behave in an abusive or corrupt manner toward them. Thus, by interacting with female officers, women's priors about the police are positively updated. Within policing, scholars have found, for

example, that when female officers are added, women are more likely to report crimes and perceptions of the police improve (Karim [2017](#); Keiser, Wilkins Meier, and Holland [2002](#); Leger [1997](#); Natarajan [2016](#)).

Additionally, social bonds may form more quickly when the police officer and the citizen share demographic characteristics. This is because it may be easier for those with shared demographic characteristics to find common ground during a social interaction. For example, women may be quicker to find common ground with other women than with men because they are more likely to share womanly experiences, such as pregnancy or childcare. Women's conversations may differ from men's, as women talk more about home and family, and use more emotive words (Haas [1979](#); Hirschman [1994](#)). Females may develop more ways to express affection and interpersonal concern in their social interactions with members of their own sex (Hall and Braunwald [1981](#)). As such, conversations between women may be less anxiety-prone and more prone to empathy formation.

In theory, any type of shared demographic characteristic could lead to improved opinions about police authority, but visible characteristics such as race, ethnicity, and sex are more likely to do so because the similarity is identifiable during a social interaction. For the purpose of this paper, I focus on sex as the shared demographic characteristic, because there is less research on the effects of the integration of women into the security forces than on the integration of different ethnic groups into the security forces.[6] This means that when female police officers interact with local women, *local women's* views about the police should improve. This leads to the following hypothesis:

Hypothesis 2. *When women living in areas of limited statehood engage in relationship-building interactions with female police officers, they are more likely to develop positive perceptions of the police than if they interact with male police officers.*

A question may arise as to whether female police officers are perceived as typical police officers. Because historically police officers have been men (including in Liberia), there is reason to believe that they are not perceived as representatives of the police. At the same time, their novelty could elicit attitudinal change among the population. However, it is not clear how attitudes of local men and women might change. On the one hand, the novelty of women in the police force could be looked on favorably, particularly by women, as mentioned above. At the same time, both men and women may have prior gender stereotypes about appropriate roles for women. As the provision of security is a predominantly masculine role, it is possible that women's inclusion may lead to resentment of the police because female inclusion violates social norms. As such, the different sets of possibilities suggest that the sign of the effect could be indeterminate. For this reason, I test the effect of female visits on local females (H2), but I also include analyses on the effect of female visits on both male and female locals as a theory-building exercise.

---

[6] On ethnic integration into police forces, see Tyler ([2005](#)).

The Procedural Justice Mechanism

In addition to the relational contact mechanism and the performance-based legitimacy mechanism, the procedural justice mechanism could also help explain how attitudes toward state agencies change as the state expands into areas of limited statehood. The mechanism posits that perceptions of fairness/unfairness during procedural interactions with state agents legitimize/delegitimize the state (Mcloughlin 2018; Sunshine and Tyler 2003; Tyler 2006). A procedural interaction is a transactional one, in which either the state is carrying out its duty or citizens are seeking goods and services from the state. Procedural interactions could take the form of experiences with the court system, court judgments, police stops, traffic tickets, or any other interactions that involve some form of procedural decision made by a state agent (Huo and Tyler 2002). It could also take the form of goods and service provision, such as medical care at a public hospital or receiving social security or an ID card from the state.

The procedural fairness mechanism could be classified as another model of relational state building. In fact, Fisk and Cherney (2017, 267) classify it as such because procedural justice requires interactions between citizens and state agents. Yet, unlike the relational contact mechanism, the goal of the interaction is not to foster a connection between two individuals; rather, the state agent is responsible for a decision that impacts the citizen's life; it has some degree of power over the citizen. Additionally, the state agent is responding to a citizen rather than seeking out the citizen to build a relationship before it provides a good. Thus, whereas the relational contact mechanism is a proactive and non-hierarchical model of relationship building, the procedural justice mechanism is a reactive, hierarchical model of relationship building. To improve opinions, the procedural justice mechanism requires state agents to be perceived as fair in their decision making during such interactions (McLoughlin 2015; Rothstein 2009; Tyler 2006). This leads to the following hypothesis:

> **Hypothesis 3.** *When citizens living in areas of limited statehood perceive police unfairness, they are less likely to develop positive perceptions of the police force.*

THE LIBERIAN CONTEXT

The three hypotheses are tested in rural Liberia. The Liberian civil wars lasted between 1980 and 2003 and claimed the lives of almost 250,000 people—mostly civilians—and led to a complete breakdown of law and order (Friedman and MacAulay 2011). It ended in a negotiated settlement, mediated in 2003, and the UN Mission in Liberia (UNMIL) was established to help keep the peace. By the end of the war, Liberia had fifteen different security agencies with overlapping functions and mandates; many police stations had been abandoned, destroyed, or taken over by rebel forces; the state lacked basic equipment, vehicles, fuel, and communications systems; and many police officers and other state officials had fled the country (Friedman and MacAulay 2011). In short, the security forces, like all other state institutions, were in dire need of rebuilding as much of Liberia lacked state authority.

At the end of the war, in conjunction with international donors, the Liberian state started to roll out a program to decentralize its state agencies, which included expanding security and rule of law into areas of limited statehood. For improving the rule of law, this meant expanding policing, particularly community policing programs and establishing a rural court system. In 2015, the LNP began recruiting more police officers to increase the number of police in rural areas. Many of these officers were trained in community policing and would be stationed as community policing officers in rural counties. It is in this context of state expansion that the field experiment below took place.

Experimental Context: Grand Kru County

The field experiment in Liberia was conducted in one of Liberia's most isolated counties: Grand Kru County.[7] Civilians who live in the county have very little interaction with the state and generally do not perceive the police favorably. The county has historically been neglected by the state, and at the end of the war, it was considered a "forgotten county."[8] In 2014, there were only 19 LNP officers assigned to the county, which had a population of 57,106 people. In 2011, only 16% of the population of the county had interacted with a police officer, which was one of the lowest percentages out of all the counties in Liberia (Vinck, Pham, and Kreutzer 2011).

Although the county has limited state presence, like other rural counties in Liberia and other areas of limited statehood, the county possesses alternative sources of governance such as chieftaincies, secret societies, and other traditional structures that compete with the state's authority.[9] Moreover, even though the UNMIL mission was present in Liberia, it had a minimal footprint in Grand Kru County. As such, Grand Kru County is representative of rural counties in which there is some minimal international presence, making it representative of most rural subunits in countries with peacekeeping missions (Ruggeri, Dorussen, and Gizelis 2018). Finally, Grand Kru's relative homogeneous ethnic balance of Kru and Grebo made it possible to implement the experiment without worrying about the possible confounding effects of ethnic heterogeneity. It is, thus, representative of other sub-Saharan countries, of which many subunits are ethnically homogeneous (Gershman and Rivera 2018). Overall, the county provided the conditions necessary for the experiment: an area of limited statehood with competing governance structures where the state is expanding and where perceptions about the police are generally negative.

---

[7] A county is a sub-national unit of governance.
[8] See: "More Aid Needed for 'Forgotten' Grand Kru County," IRIN News, 2005, http://www.irinnews.org/report/54819/liberia-more-aid-needed-forgotten-grand-kru-county.
[9] See Ellis (2007).

The Ethics of Doing Fieldwork (with Police Officers) in Areas of Limited Statehood

Lake and Cronin-Furman ([2018](#)) have recently called on scholars who engage in fieldwork to incorporate a section on ethics in the research design section of papers. Although IRB approval is the minimal requirement necessary to meet ethical standards,[10] this field experiment necessitated more careful ethical considerations because of the nature of the manipulation, the research context, and safety concerns of the enumerators. First, it would have been unethical to introduce state expansion through the field experiment into a county and then discontinued the services after the implementation of the experiment, as it would have raised expectations about the police that the state could never meet. To remedy this, the experiment was rolled out just before the LNP's plan to expand police into the rural counties began. Moreover, the field experiment required my research team (including police officers) to spend three months in the field, often walking long distances and potentially getting sick when they would be far away from a health clinic. Thus, as a team, we created a safety plan and determined the conditions under which the research would be delayed or stopped. Finally, it was important to ensure that introducing law enforcement to communities (that have little prior interactions with them) would not bring harm to them or to the police officers. This was important for the experiment itself as one of the mechanisms required that the police behave in a non-predatory manner. The police officers in my team were trained in community policing techniques. Additionally, one enumerator was assigned to each police team, which meant that they monitored not only the correct implementation of the experiment but also the behavior of the police. Although the enumerators' presence perhaps compromised external validity, in this case it was necessary to ensure that officers did not harm citizens and that citizens did not harm police officers.

RESEARCH DESIGN

Treatments

The above sections provided three different mechanisms for how expansion of police presence into areas of limited statehood might affect perceptions of the police. I experimentally test the relational contact mechanism (H1 and H2) and observationally test the procedural justice mechanism (H3). The treatment for the relational contact mechanism was a 20–30 minute, face-to-face household visit by police officers. The treatment used to test H2 was a 20–30 minute, face-to-face household visit by female police officers compared with a 20–30 minute, face-to-face household visit by male police officers. In total, two male officers and two

female officers (four total) were selected for the study and memorized a scripted visit, which ensured that the household visits were similar in content.[11] Officers were required to wear their uniform and introduce themselves as the LNP, so that there was no confusion about who they represented. Mixed-sex teams were not possible for two reasons. First, testing H2 required all female (and all male) teams of police officers. Second, there were insufficient funds to add a third treatment arm.

Officers for the project were chosen based on similarity in their personalities. This was done to ensure that specific personality traits would not drive the treatment effect. The LNP in Monrovia provided 15 Kru/Grebo-speaking officers for the study. I then filmed each of the 15 officers giving the speech to an audience, as if they were engaging in community policing. The videos were then shown to a group of Kru/Grebo people in Monrovia who rated the officers on different characteristics.[12] Four officers were chosen (two female and two male) based on similarity in personality scores both across and within teams. The officers also engaged in one month of intensive training, including a pilot test in a different county, to ensure that they behaved the same way during the visits.

At the end of the visit, the officers gave the household cards with phone numbers of the local police. These cards were later used to identify whether individuals within the households had been treated. The officers spent between one and three continuous days in each village and were required to remain in their designated neighborhoods during the entire time.

Experimental Design

The research design used an intent-to-treat design. When the police officers visited the households, they treated the entire household. That is, they spoke to and interacted with all the members of the household who were present at the time of the visit. Individuals randomly selected for the endline survey were 18 and over.[13]
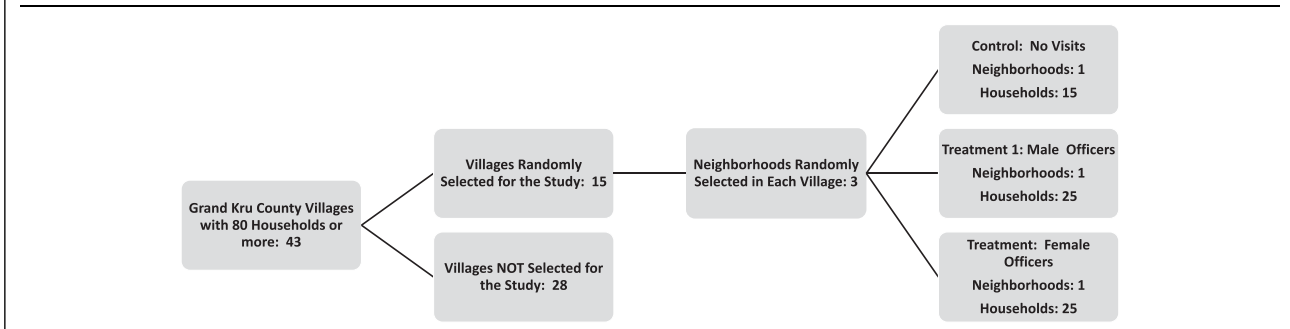
According to the 2008 census, 43 villages included 80 households or more. From these, 15 were randomly selected for the study and constitute a representative sample of the county. The Online Appendix provides

---

[10] All parts of the study underwent an Internal Review Board process. On May 12, 2015, the Emory University's IRB approved the project. The IRB is filed under IRB00073869.

[11] The LNP script is found in the Online Appendix, Section 12. It was written by the Deputy Director of The Police Training Academy based on what police officers should say when they go into the field.

[12] After watching the video of each officer, they were asked: Do you want this officer to be your friend? Do you think the officer knows his or her work good? Are you afraid of the officer? Does the officer make you feel tired? Would you follow the officer's instructions? Do you like this officer? Would you talk to this officer if they came to your home? Would you feel fine asking the officer questions? Would you feel fine reporting a crime to the officer? Should the officer be sent to Kru/Grebo-speaking communities to promote the LNP? List some words to describe the officer. The suggestions for the questions and their wording came from the Liberian enumeration team.

[13] Of survey respondents who stated that they did not meet with police officers during the visit, 47% were women and 53% were men, indicating that there was no bias in who the police officers interacted with during their visits.

FIGURE 1. Research Design

TABLE 1. Research Design for Treatment

Group	Number of villages	Households per village	Total households
Treatment 1: visit by male police	15	25	375
Treatment 2: visit by female police	15	25	375
Control: no visit	15	15	225
Total	15	65	975

descriptive comparisons of the villages selected and not selected for the study. Within each village, a randomly selected number of houses were visited by female, male, and no police officers (see Figure 1). Within each village, randomization occurred at the neighborhood level. When the team, which included four police officers and three enumerators, arrived in the villages, they asked for a list of the neighborhoods in the village. On average, in the 15 villages selected for treatment, each village included four neighborhoods, and each neighborhood contained about 56 households. Neighborhoods were randomly assigned to the female police officer treatment, male police officer treatment, and the control group. The target for each treatment group was to reach 25 households. They used a random walk technique to randomly choose houses in the randomly assigned neighborhood. The total sample size consisted of 750 households, 375 visits by male officers, and 375 visits by female officers. Table 1 provides the research design for the experiment.

Outcomes

Three weeks after the officers' visits, the Center for Applied Research and Training (CART) surveyed the treated households. The survey included questions that ranged from maternal healthcare to security. The enumerators presented themselves as an independent agency doing a survey on health and security in the county. The survey was conducted three weeks after the treatment for both theoretical and practical reasons. Theoretically, because there were no systematic repeat visits, the treatment effect was likely to decay quickly. The goal was not to capture decay or longer-term effects but rather to better understand whether opinions changed at all in the short term when exposed to a first-

time relationship-building encounter. Second, the surveys needed to be completed before the rainy season, as it would have been too dangerous for the enumerators to travel during that time. The enumeration team surveyed the exact same households as those that were visited by police officers. They were able to recognize the households based on maps that were drawn by the enumerators who were a part of the police team, and by a code that was written discreetly adjacent to the household. All treated households were located.

To avoid spillover, the control group survey was conducted at the same time as the treatment was implemented, which minimized the number of treated household members communicating to those in the control group about the police visits. The time lag in survey implementation, however, presents several problems. First, it may have caused a different group of people to be available for the survey due to farming and travel schedules. However, the enumerators used the same sampling method for both sets of surveys to ensure that all household members had an equal chance of being selected for the survey. Second, individuals who were surveyed earlier may have told other members of the community about a survey, thereby priming them about the survey prior to the implementation of the treatment group survey. To account for this suspicion, the Online Appendix provides models that include those who were suspicious about the survey as a control as well as a control for whether other people were nearby at the time of the survey.

Specific to this study and consistent with a dyadic and monadic understanding of legitimacy, I assessed two types of questions: preferences for the police over non-state and external actors (dyadic) and positive perceptions of the police (monadic). To address the former, I included

questions about preferences for the police to respond to two security crisis situations: a community dispute (hala hala) and mob violence. I coded whether respondents preferred the police over UNMIL, community leaders, community watch teams, traditional leaders, or others.[14]

Prefer Police (for Security Response)

- If there is a hala hala [dispute] between two tribes or religions in your community, who would you most like to resolve the situation?
- If there is a big group of people making noise to fight [mob violence], who would you most like to resolve the situation?

To measure positive perceptual outcomes, the survey included questions about perceptions of police restraint (abuse) and police effectiveness. The questions were asked in statement form and respondents were asked whether they strongly agree, agree, disagree, or strongly disagree.[15] I grouped the "strongly disagree" and "disagree" answers together (coded as 0) and the "strongly agree" and "agree" together (and coded as 1).[16] The Online Appendix shows the correlation coefficients between the questions in each set. There is variation in item intercorrelations within question sets, especially the questions for perceptions of effectiveness. However, the pre-analysis plan specified these groupings and the correlation matrices in the Online Appendix show that the two sets of questions are not highly correlated with one another, indicating that they are capturing different concepts (abuse and effectiveness).

Police Abuse

- The police can sometimes steal things from me
- The LNP are causing problems (shouting, yelling, etc.) when they come into the community
- The LNP sometimes behave like criminals
- The LNP discriminate based on religion/ethnicity/tribe
- The LNP are corrupt and eating money

Police Effectiveness

- No rogue [criminal] will come into the village if the LNP stay here
- There will be less crime in the community if the LNP come to the village
- The LNP are good at catching criminals
- I feel safer when the LNP are in the village
- I am afraid that the LNP will arrest me some day if I commit a crime.

For each set of questions, I create an additive index. The index assumes that each question is equivalent to the other questions in measuring the overall concept.

Spillover and H3

In addition to lagging the time of the surveys, to prevent spillover, the treatment was assigned at the neighborhood level. Neighborhoods are generally far apart in rural villages, which means that those in a control neighborhood would not have interacted with police officers in a treated neighborhood. The police officers also spent the night in their designated neighborhood, so that they would not be seen in other parts of the village at night.

Despite these precautions, there is evidence of a violation of the stable unit treatment value assumption (SUTVA). Specifically, 45 individuals in the control group had knowledge about the police visits. The spillover group, however, allowed me to observationally test H3—the procedural justice mechanism. The procedural justice mechanism requires the police to make decisions about procedural services and requires that the public perceive the procedure in a fair way. This did not occur with the randomization process. The spillover group felt they did not get equitable treatment from the police. They felt that they were treated unfairly because they did not get a household visit; they felt as if they were missing out on an important government service.[17]

Models

In addition to presenting the results in a regression framework, I present the Average Effect Size (AES), which measures across all questions within each cluster (prefer police, abuse, and effectiveness), following the procedure proposed in Clingingsmith, Khwaja, and Kremer (2009) and Kling, Liebman, and Katz (2007). The AES across J related dependent variables is given by $\tau = \sum_{j=1}^{J} \frac{\pi_j}{\sigma_j}$, where π_j is the average treatment effect on each dependent variable and σ_j is the standard deviation of dependent variable j in the control group. To test the null hypothesis of no average effect, the effects π_j are jointly estimated using a seemingly unrelated regression. The J dependent variables are stacked to compute a variance–covariance matrix for testing the statistical significance of τ, the AES.[18] AES coefficients are interpreted in terms of standard deviations from the control group mean. All models are presented with standard errors clustered at the neighborhood level. In addition to AES, I present a logit model to test Hypothesis 3, which treats the data as observational.

Response Bias and Internal Validity

It is possible that respondents may have felt the need to say positive things about the police out of fear of their

---

[14] These questions were pre-specified in the pre-analysis plan. See Online Appendix for distribution of answers.
[15] All models below drop "I refuse to answer" (of which there were very few) and "I don't know" answers. See Online Appendix for analysis of these responses.
[16] The results do not change if ordinal outcomes are used. All questions were asked in and are presented in Liberian English.

[17] One of my research assistants returned to three villages to speak to a dozen of these individuals.
[18] For further details, see Clingingsmith, Khwaja, and Kremer (2009) and Kling, Liebman, and Katz (2007).

return. The enumerators collected data on the reactions of the participants at the beginning, during, and at the end of the visit. On average, the enumerators recorded that about 24% of the participants were anxious at the beginning of the visit, but that 99% of the participants were appreciative of the visit after it was over and that 96% wanted the police to return to their house in the future. This suggests that by the end of the visit, participants in the study were not scared of the police and that they likely did not answer questions on the survey due to fear of the police. Furthermore, the interactions with the police decreased anxiety about the police over the course of the visits, providing evidence of the social bond mechanism that is necessary for relationship building.

Additionally, to verify that the relational contact mechanism contributed to improved perceptions and not the content of the speech, which included information about police professionalization, I conducted a survey experiment in a different set of villages in Grand Kru County. Enumerators read the same script to randomly selected citizens during the implementation of a survey and measured the same responses. The results show that this information actually led to a decrease in preferring the police to respond to security crises and a decrease in perceptions of police effectiveness (see Online Appendix). There was a small decrease in perceptions of abuse but not to the same degree as the police visits. These findings confirm that it was not the content of the speech, but rather the relational dynamics of the visit that helped change perceptions.

Randomization Checks and Other Threats to Inference

The data for the randomization checks come from the surveys implemented posttreatment. A baseline survey was not possible because it would have aroused suspicion among community members.[19] Members of communities are not used to receiving visits by strangers. About 62% had never interacted with community outsiders in their own village. Due to the novelty of strangers visiting villages, the study took great care to disassociate the police visits from the survey—adding another survey would have made it difficult to decouple the surveys from the police visits.

Despite randomizing treatment at the neighborhood level, balance was mixed (see Online Appendix). Compared with those that did not experience visits by police, those that experienced visits by female officers tended to be less Christian, own land, farm, and have a higher number of household members. Balance was better between those that experienced visits by male officers and the control group. Those that experienced visits by male officers were less likely to be traditional leaders and more likely to have a larger household. Imbalance may be because the randomization was done

at the neighborhood level. Some neighborhoods may have had higher numbers of land owners, farmers, Christians, larger families, or traditional leaders. However, the imbalanced covariates are unlikely to influence perceptions of the state security forces, as they are not common variables in predicting police perceptions. Nevertheless, they are included in the models below.

Finally, there may be an issue of multiple testing as I include multiple dependent variables, two treatments, and assess the heterogeneous treatment effects of respondents' sex. To account for this, I provide adjusted p values in the Online Appendix. I also present AES of the dependent variables, which mitigates Type I and II errors (Clingingsmith, Khwaja, and Kremer 2009; Kling, Liebman, and Katz 2007).

RESULTS

Hypothesis 1 stated that relationship-building interactions with police officers should lead to locals developing positive perceptions of the police. First, I observe the results for the dyadic indicators of legitimacy. Model 1 (Table 2) shows that the pooled treatment effect (visits by male and female police officers) on preferences for the police to respond to crisis situations is positive and statistically significant. On average, those in the control group responded that they wanted the police to respond to one security crisis, but those who were treated wanted the police to respond to both security situations. This means that the treatment increased citizens' preferences for the police over alternative sources of security for a wider range of security issues.

Table 3 provides the AES estimates of police visits of the pooled treatment. The coefficients are interpreted in terms of standard deviations from the control group mean. Because all the component dependent variables are binary, the coefficients are interpreted in terms of percentage point differences between the treatment and control groups. For police response to crises, the pooled treatment increased the control group mean by 1.49 standard deviations. The component dependent variables show a 70 point increase for the police to respond to a hala hala and a 61 point increase for the police to respond to mob justice. Thus, not only did citizens prefer the police for a wider range of security issues, but also there was a relatively large shift in these preferences. This could be due to social bonds created with the new state security actor. If citizens began to trust the individual police officers who visited them, then they may prefer those specific officers to come back and respond if there is a dispute in the community. The overall indication is that even a one-time visit with a police officer could lead to switches in preference for state security over informal/external sources of security.

Next, I assess the results of the monadic indicators of legitimacy. Model 2 (Table 2) shows that visits by police officers led to a negative and statistically significant effect for perceptions of police abuse. On average, respondents in the control group agreed with three

---

[19] About 10% of the respondents were suspicious about the survey being related to the police visit.

TABLE 2. Pooled and Individual Treatment Effects

	(1) Prefer police	(2) Abusive	(3) Effective	(4) Prefer police	(5) Abusive	(6) Effective	(7) Prefer police	(8) Abusive	(9) Effective
Pooled treatment	1.23***	−1.99***	0.93***						
	(0.14)	(0.34)	(0.18)						
Traditional leader	−0.05	0.29[+]	0.22	−0.05	0.29[+]	0.23	−0.07	0.36[+]	0.20
	(0.09)	(0.17)	(0.14)	(0.10)	(0.17)	(0.14)	(0.12)	(0.19)	(0.13)
Land	−0.03	−0.18	0.14	−0.03	−0.18	0.15[+]	−0.08	−0.10	0.07
	(0.06)	(0.11)	(0.09)	(0.06)	(0.12)	(0.09)	(0.06)	(0.12)	(0.09)
Christian	0.39**	−0.35	0.29	0.39**	−0.34	0.25	0.41**	−0.34	0.36
	(0.15)	(0.34)	(0.23)	(0.14)	(0.34)	(0.22)	(0.15)	(0.34)	(0.24)
Household number	−0.009	−0.008	−0.002	−0.009	−0.008	−0.001	−0.007	0.003	−0.02
	(0.008)	(0.02)	(0.02)	(0.008)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
Male police treatment				1.23***	−2.00***	0.98***			
				(0.14)	(0.34)	(0.20)			
Female police treatment				1.23***	−1.99***	0.87***			
				(0.15)	(0.36)	(0.19)			
Female police treatment (compared with male police treatment)							0.005	−0.02	−0.08
							(0.06)	(0.14)	(0.14)
Constant	0.23	3.62***	2.60***	0.23	3.62***	2.63***	1.45***	1.50***	3.65***
	(0.20)	(0.47)	(0.27)	(0.20)	(0.47)	(0.27)	(0.17)	(0.38)	(0.27)
N	966	833	883	966	833	883	744	675	681
R^2	0.37	0.25	0.11	0.37	0.25	0.11	0.02	0.009	0.009
BIC	2,010.4	2,976.4	2,782.5	2,017.3	2,983.1	2,787.7	1,451.6	2,325.4	2,135.8

Standard errors are in parentheses.
[+] $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

TABLE 3. Average Effect Size

	(1) Police to respond to crisis	(2) Perceptions of abuse	(3) Perceptions of effectiveness
Pooled police treatment	**1.49 (0.16)*****	**−0.70 (0.11)*****	**0.36 (0.10)*****
Hala hala	0.70 (0.07)***		
Mob justice	0.61 (0.08)***		
Steal		−0.64 (0.10)***	
Cause problems		−0.26 (0.06)***	
Like criminal		−0.44 (0.09)***	
Corrupt		−0.26 (0.07)***	
Discriminate		−0.51 (0.08) ***	
No rogue			0.04 (0.10)
Less crime			0.10 (0.03)**
Catch criminal			0.31 (0.08)***
Feel safer			0.11 (0.06)+
Arrest me			0.36 (0.07)***
N	969	969	970

Notes: AES for each cluster of outcomes.
Standard errors are in parentheses.
AES coefficients are interpreted in terms of standard deviations from the control group mean.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

questions about police abuse, but the treatment reduced that number by two questions. Additionally, Model 3 shows that the visits led to a positive and statistically significant effect for perceptions of police effectiveness. On average, those in the control group agreed with three questions about police effectiveness, but the treatment increased the number by almost one question. This means that the police visits increased the range of positive perceptions about the police.

Looking at the AES estimates in Table 3, for perceptions of police abuse, the pooled treatment decreased the control group mean by 0.70 standard deviations. The component dependent variables show a 64 point decrease in believing the police will steal, a 26 point decrease in believing the police will cause problems in the community, a 44 point decrease in believing the police behave like criminals, a 26 point decrease in perceiving the police as corrupt, and a 51 point decrease in perceiving the police as discriminatory. For perceptions of effectiveness, the pooled treatment increased the control group mean by 0.36 standard deviations. The component dependent variables show a 10 point increase in believing police presence will lead to less crime, a 31 point increase in believing the police can catch criminals, an 11 point increase in feeling safer, and a 36 point increase in believing that the police will arrest them if they commit a crime.

The results indicate that the police visits shifted attitudes about the police, particularly with respect to police abuse. In particular, they show that the interactions led to positive updating of negative priors—if police abuse constituted citizens' priors, then the police's non-abusive behavior led to positive updating of citizens' priors. Although the police demonstrated that they are not abusive (e.g., by not being abusive during their visits), they did not necessarily show that they are effective (e.g., they did not apprehend

a criminal), which is perhaps the reason for the weaker effects (half the AES for "abusive") on perceptions of effectiveness. Nevertheless, once again the overall indication is that even a one-time visit with a police officer, in which the police officer behaves in a non-abusive way, could lead to relatively large increases in positive perceptions of the police.

As mentioned above, the study suffered from spillover. The spillover occurred nonrandomly, as those who knew about the police visits were younger and more likely to be landowners. The inclusion of the spillover group in the control group created a larger treatment effect than would otherwise be the case, because knowledge about the police visits among control group respondents led to the development of negative opinions about the police. To account for this bias, I excluded the spillover group from the analysis. This yielded very similar results to those presented above (see Online Appendix). Moreover, the results are robust to the use of Complier Average Causal Effects (CACE), which accounts for noncompliance (see Online Appendix).

In the Online Appendix, I address other potential threats to inference, including multiple testing, omitted variable biases, and dropped responses (those who answered "I don't know"). Support for H1 remains strong after these robustness checks.

Hypothesis 2 suggested that relational bonds should be stronger among those who share demographic characteristics. Table 4 shows that there was no support for this hypothesis. When women were visited by female police officers, there was no *added* improvement in perceptions of the police among local females.[20] This means that women viewed female police officers in the

---

[20] Similarly, local men's opinions about the police did not improve after visits by male police officers (see Online Appendix).

TABLE 4. Heterogeneous Treatment Effect: Local Females Treated by Female Police Officers

	(1) Prefer police	(2) Abusive	(3) Effective
Female police treatment	0.004	−0.05	−0.11
	(0.08)	(0.18)	(0.15)
Female	0.13[+]	−0.31*	0.10
	(0.07)	(0.14)	(0.14)
Female police treatment × female	−0.01	0.11	0.04
	(0.10)	(0.19)	(0.17)
Traditional leader	−0.03	0.29	0.23[+]
	(0.12)	(0.19)	(0.14)
Land	−0.08	−0.10	0.07
	(0.06)	(0.12)	(0.10)
Christian	0.39*	−0.31	0.35
	(0.16)	(0.35)	(0.24)
Household number	−0.007	0.003	−0.02
	(0.01)	(0.02)	(0.02)
Constant	1.40***	1.62***	3.61***
	(0.19)	(0.41)	(0.28)
N	744	675	681
R^2	0.03	0.02	0.01
BIC	1,457.1	2,331.7	2,147.2

Standard errors are in parentheses.
[+] $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

same way as male officers. They may not necessarily have found it easier to bond with female police officers. Models 4–6 (Table 2) show the treatment effect based on if the visits were conducted by male or female police officers, and Models 7–9 compare female police officers' visits with male police officers' visits. These tables show that male and female police officers equally improved police perceptions during their household visits.[21]

The null results are perhaps due to the police officers' similarity in personality. Choosing police officers with the same personality meant that the social interactions were likely to be similar as well. Thus, that there are no differences between the visits by male and female police officers is consistent with the idea that the social interactions are driving the results. It also means that the treatment to test H2 was weak. The assumption behind H2 was that similarity in sex should have been enough to ensure common ground among women because sex is a visible signal of similarity. At the same time, women are not always drawn to other women as the category of "woman" is diverse, and women's solidarity is predicated on much more than similarity in sex.[22] Moreover, the female police officers chosen for this study may not have elicited social bonds because they did not necessarily behave in a feminine way nor ascribe to their gender role. The selection of officers biased against selecting feminine women, because the female police officers chosen had similar personalities to men that were chosen. Research finds that women in the security forces tend to take on masculine traits to better fit into

the organization (Rabe-Hemp 2009). Thus, in this study, the female officers likely behaved like men and subsequently elicited the same reaction as men behaving like men, but because the female officers behaved like men, the female officers did not elicit stronger perceptional change among women.

Despite the null findings, the results of H1 and H2, taken together, are encouraging. For one, it means that female police officers elicit the same response among local men and women as male police officers. It shows that local men and women are not put off by female police officers. This should assuage some policy makers' fears that female police officers are not as competent as male police officers.[23] For the purposes of theory building, it suggests that the personality of police officers, including their gendered performances, is perhaps more important for policing than police officers' sex.

I now turn to Hypothesis 3. Hypothesis 3 stated that when citizens living in areas of limited statehood perceive police unfairness, they are less likely to develop positive perceptions of the police. To test H3, I draw on the spillover group from above. Recall that those who found out about the police officers' visits but who were not selected to be visited found this exclusion to be unfair. They felt that they were missing out on an important police service. The procedural decision in this case was the decision to not visit all households in the village, and this decision resulted in perceptions of unfairness. Thus, the unintended consequence from random selection allowed me to test H3 in an observational manner. Even though the test should not be interpreted as causal, it is perhaps an improvement over

---

[21] The AES estimates for these models are in the Online Appendix.
[22] See, for example, discussions about intersectionality: Crenshaw (1990) and Hooks (1981).

[23] See, for example (Bacon 2017, 380).

TABLE 5. Perceive Police as Unfair

	(1) Prefer police	(2) Abusive	(3) Effective
Knew about police visits (perceived police as unfair)	−0.61***	1.42**	−0.44*
	(0.13)	(0.35)	(0.16)
Age	0.001	0.001	0.007
	(0.003)	(0.01)	(0.005)
Land	0.13	−0.56	0.41*
	(0.17)	(0.37)	(0.18)
Constant	0.55*	3.19***	2.64***
	(0.21)	(0.44)	(0.21)
N	221	157	201
R^2	0.11	0.10	0.07
BIC	524.9	627.6	649.9

Standard errors are in parentheses.
$^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

other tests of the procedural justice mechanism that have relied on self-reports (Huo and Tyler 2002).

Table 5 compares the spillover group with the remainder of the control group and shows that those who knew about the police visits were much less likely to prefer the police to respond to crisis situations, they were more likely to perceive the police as abusive, and they were less likely to perceive the police as effective.[24] Substantively, those in the control group preferred the police to respond to one type of security threat, but those in the spillover group did not want the police to respond to a single security threat. Those in the control group agreed with two of the five questions about police abuse, but those in the spillover group agreed with four of the five questions about police abuse. The findings indicate that when the police make decisions that are perceived as unfair, the decisions have much larger consequences on the perceptions of legitimacy of the police as a whole.

The unintended consequence from this experiment draws attention to the ethics of randomizing goods and service delivery. To address the ethical issue posed by randomization in this case, I ensured that the LNP revisited the villages in our study as a part of their community policing expansion. From a broader experimental point of view, however, this finding has implications for randomized controlled trials that assume that randomization will be perceived as fair. Spillover may not only contaminate causal effect estimates but may also have corrosive effects on outcomes of interest. Here, even though randomization (called "lucky ticket" in Liberia) was explained to community elders, the explanation did not trickle down to citizens. This means that future studies should take care to better understand how information about randomization is spread and interpreted. They should be particularly aware of control group perceptions about randomization. There should also be mechanisms in place to account for any adverse consequences from perceived unfairness.

---

[24] The results here are robust to the inclusion of the treatment groups in the model (see Online Appendix).

DISCUSSION AND CONCLUSION

As states expand their authority into areas of limited statehood, under what conditions does this state building improve perceptions of state authority? This paper suggests that, in addition to institutional or identity-based sources of legitimacy, relational state building, whether preemptive (relational contact) or reactive (procedural justice), is important for building legitimacy in state institutions. Using a field experiment in rural Liberia, I show that there is support for both mechanisms. When female and male police officers engaged in face-to-face household visits with locals, locals were more likely to prefer the police to respond to security crisis, less likely to perceive the police as abusive, and more likely to perceive the police as effective. The findings demonstrate that Risse and Stollenwerk's (2018) conceptualization of social trust and esteem as a source of legitimacy apply to state agents as well as with non-state actors. More importantly, it means that trust in authority can be engineered through relationship-building activities with agents of the state.

The results of the study also have particular implications for contact theory. They broaden the context in which contact theory can be applied to areas of limited statehood, and they increase the types of actors to which the theory can be applied to citizens and state agents. At the same time, they show that two of the mechanisms through which contact theory operates—(1) positive updating of negative priors and (2) social bonds (Pettigrew 1998)—are potential pathways for positive opinion formation.

The study found that shared demographic characteristics between citizens and the police did not strengthen the relational contact mechanism, as female police officers were no more likely than male officers to elicit positive perceptions among local women. This null result perhaps shed light on the limitations of passive representation, the theory used to develop H2 (Mosher 1982, 12). In particular, participation in groups that have strong identities of their own, such as policing, could overshadow demographic-based identities. The evidence

from the experiment points in this direction—local women may have seen the female police officers as "police officers" before they saw them as "women." Future studies could test the effect of a possibly stronger "gender treatment," which would test the effects of masculine versus feminine police officers on opinion formation.

Finally, the paper tests the procedural justice mechanism and finds support for it. Locals who believed that they missed out on household visits by the police developed negative opinions about the police. The unintended consequences of the experiment allowed for a novel way to test the procedural justice mechanism as most research on it relies on self-reports (Huo and Tyler 2002). Support for the procedural justice mechanism using a different way to test it suggests that the theory is robust.

Although the paper has found support for relational state building, it may not lead to positive opinion formation in all contexts. Because information updating is a key component of the relational contact mechanism, the theory may not apply (or be weaker) in areas where priors are more positive about the state. The effects of the relationship-building activity may also decay over time unless that relationship is sustained. It is not clear if relationship building elicits trust for all governance actors. Future studies could thus explore the degree to which citizens' priors affect opinion formation, whether relationship building over the long term leads to sustained effects, whether other state actors (other than police) or even non-state actors (e.g., non-state—rebel groups or external actors—peacekeepers, NGOs) benefit from relationship building, and whether other shared demographic characteristics (e.g., gender, race, or ethnicity) strengthen social bonds.

Studies such as this one become increasingly important as international actors, such as peacekeepers, continue to press weak states to expand their authority, particularly their security presence, into areas of limited statehood. This paper has shown that there may not be drawbacks from this expansion if state builders take care to ensure that state agents engage in relationship-building activities with locals, and if they ensure that procedural interactions between state agents and citizens are perceived as fair. As such, even though the enforcement arm of the state—the police—may have been predatory at one point, their expansion into new territories perhaps gives them an opportunity to be sources of legitimacy in the future.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit https://doi.org/10.1017/S0003055419000716.

Replication materials can be found on Dataverse at: https://doi.org/10.7910/DVN/OARGFC.

REFERENCES

Allport, Gordon W. 1954. *The Nature of Prejudice.* Oxford, England: Addison-Wesley.

Arjona, Ana, Nelson Kasfir, and Zachariah Mampilly. 2015. *Rebel Governance in Civil War.* Cambridge, UK: Cambridge University Press.

Bacon, Laura. 2017. "Liberia's Gender-Sensitive Police Reform: Improving Representation and Responsiveness in a Post-Conflict Setting." *International Peacekeeping* 22 (4): 372–97.

Beath, Andrew, Fotini Christia, and Ruben Enikolopov. 2013. "Empowering Women through Development Aid: Evidence from a Field Experiment in Afghanistan." *American Political Science Review* 107 (3): 540–57.

Blair, Robert A., Sabrina M. Karim, and Benjamin S. Morse. 2019. "Establishing the rule of law in weak and war-torn states: Evidence from a field experiment with the Liberian National Police." *American Political Science Review* 113 (3): 641–57.

Braithwaite, Valerie, and Margaret Levi. 2003. *Trust and Governance.* New York City: Russell Sage Foundation.

Brass, Jennifer N. 2016. *Allies or Adversaries: NGOs and the State in Africa.* Cambridge, UK: Cambridge University Press.

Broockman, David, and Joshua Kalla. 2016. "Durably Reducing Transphobia: A Field Experiment on Door-to-Door Canvassing." *Science* 352 (6282): 220–4.

Chong, Dennis, and James N. Druckman. 2010. "Dynamic Public Opinion: Communication Effects over Time." *American Political Science Review* 104 (4): 663–80.

Chong, Dennis, and James N. Druckman. 2012. "Counterframing Effects." *The Journal of Politics* 75 (1): 1–16.

Clingingsmith, David, Asim Ijaz Khwaja, and Michael Kremer. 2009. "Estimating the Impact of the Hajj: Religion and Tolerance in Islam's Global Gathering." *The Quarterly Journal of Economics* 124 (3): 1133–70.

Crenshaw, Kimberle. 1990. "Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color." *Stanford Law Review* 43: 1241.

Ellis, Stephen. 2007. *The Mask of Anarchy: The Destruction of Liberia and the Religious Dimension of an African Civil War.* New York City: NYU Press.

Fisk, Kylie, and Adrian Cherney. 2017. "Pathways to Institutional Legitimacy in Post-Conflict Societies: Perceptions of Process and Performance in Nepal." *Governance* 30 (2): 263–81.

Friedman, Jonathan, and Christine MacAulay. 2011. "Building Civilian Police Capacity: Post-Conflict Liberia 2003–2011." In *Innovations for Successful Societies.* Princeton, NJ: Princeton University.

Gershman, Boris, and Diego Rivera. 2018. "Subnational Diversity in Sub-Saharan Africa: Insights from a New Dataset." *Journal of Development Economics* 133: 231–63.

Greenhill, Kelly M., and Ben Oppenheim. 2017. "Rumor Has It: The Adoption of Unverified Information in Conflict Zones." *International Studies Quarterly* 61 (3): 660–76.

Haas, Adelaide. 1979. "Male and Female Spoken Language Differences: Stereotypes and Evidence." *Psychological Bulletin* 86 (3): 616.

Hall, Judith A., and Karen G. Braunwald. 1981. "Gender Cues in Conversations." *Journal of Personality and Social Psychology* 40 (1): 99.

Hirschman, Lynette. 1994. "Female–Male Differences in Conversational Interaction." *Language in Society* 23 (3): 427–42.

Hooks, Bell. 1981. "Ain't I a Woman Black Women and Feminism." Boston: South End Press.

Huo, Yuen J., and Tom R. Tyler. 2002. *Trust in the Law: Encouraging Public Cooperation with the Police and Courts.* New York: Russell Sage Foundation.

Joseph, Nathan, and Nicholas Alex. 1972. "The Uniform: A Sociological Perspective." *American Journal of Sociology* 77 (4): 719–30.

Karim, Sabrina. 2017. "Restoring Confidence in Post-Conflict Security Sectors: Survey Evidence from Liberia on Female Ratio Balancing Reforms." *British Journal of Political Science* 49 (3): 799–821.

Keiser, Lael R., Vicky M. Wilkins, Kenneth J. Meier, and Catherine A. Holland. 2002. "Lipstick and Logarithms: Gender, Institutional Context, and Representative Bureaucracy." *American Political Science Review* 96 (3): 553–64.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1): 83–119.

Krasner, Stephen D. 1999. *Sovereignty: Organized Hypocrisy.* Princeton, NJ: Princeton University Press.

Krasner, Stephen D., and Thomas Risse. 2014. "External Actors, State-Building, and Service Provision in Areas of Limited Statehood: Introduction." *Governance* 27 (4): 545–67.

Lake, David A. 2016. *The Statebuilder's Dilemma: On the Limits of Foreign Intervention*. Ithaca: Cornell University Press.

Lake, Milli, and Kate Cronin-Furman. 2018. "Ethics Abroad: Research and Field Methods in Areas of Limited Statehood." *Political Science and Politics*. 51 (3): 607–14.

Leger, Kristen. 1997. "Public Perceptions of Female Police Officers on Patrol." *American Journal of Criminal Justice* 21 (2): 231–49.

Levi, Margaret, and Audrey Sacks. 2009. "Legitimating Beliefs: Sources and Indicators." *Regulation & Governance* 3 (4): 311–33.

Levi, Margaret, Audrey Sacks, and Tom Tyler. 2009. "Conceptualizing Legitimacy, Measuring Legitimating Beliefs." *American Behavioral Scientist* 53 (3): 354–75.

Mann, Michael. 1984. "The Autonomous Power of the State: Its Origins, Mechanisms and Results." *European Journal of Sociology/ Archives européennes de sociologie* 25 (2): 185–213.

McLoughlin, Claire. 2015. "When Does Service Delivery Improve the Legitimacy of a Fragile or Conflict-Affected State?" *Governance* 28 (3): 341–56.

Mcloughlin, Claire. 2018. "When the Virtuous Circle Unravels: Unfair Service Provision and State De-Legitimation in Divided Societies." *Journal of Intervention and Statebuilding* 12 (4): 527–44.

Meier, Kenneth John. 1975. "Representative Bureaucracy: An Empirical Analysis." *American Political Science Review* 69 (2): 526–42.

Meier, Kenneth J., and Jill Nicholson-Crotty. 2006. "Gender, Representative Bureaucracy, and Law Enforcement: The Case of Sexual Assault." *Public Administration Review* 66 (6): 850–60.

Mosher, Frederick C. 1982. *Democracy and the Public Service*. Oxford, UK: Oxford University Press on Demand.

Natarajan, Mangai. 2016. *Women Police in a Changing Society: Back Door to Equality*. London: Routledge.

Nussio, Enzo, Miguel Garcia-Sanchez, Ben Oppenheim, and Sebastian Pantoja-Barrios. 2019. "Testing Statebuilding's 'Missing Link': Effects of Government Communication in Colombia." *Journal of Development Studies*, 1–18. Published online 18 March 2019.

Paluck, Elizabeth Levy. 2009. "Reducing Intergroup Prejudice and Conflict Using the Media: A Field Experiment in Rwanda." *Journal of Personality and Social Psychology* 96 (3): 574.

Paris, Roland. 2004. *At War's End: Building Peace after Civil Conflict*. Cambridge: Cambridge University Press.

Pettigrew, Thomas F. 1998. "Intergroup Contact Theory." *Annual Review of Psychology* 49 (1): 65–85.

Pettigrew, Thomas F., and Linda R. Tropp. 2006. "A Meta-Analytic Test of Intergroup Contact Theory." *Journal of Personality and Social Psychology* 90 (5): 751.

Pettigrew, Thomas F., Linda R. Tropp, Ulrich Wagner, and Oliver Christ. 2011. "Recent Advances in Intergroup Contact Theory." *International Journal of Intercultural Relations* 35 (3): 271–80.

Pitkin, Hanna F. 1967. *The Concept of Representation*. Berkeley, CA: University of California Press.

Rabe-Hemp, Cara E. 2009. "POLICEwomen or policeWOMEN? Doing Gender and Police Work." *Feminist Criminology* 4 (2): 114–29.

Risse, Thomas. 2011. *Governance without a State? Policies and Politics in Areas of Limited Statehood*. New York: Columbia University Press.

Risse, Thomas, and Eric Stollenwerk. 2018. "Legitimacy in Areas of Limited Statehood." *Annual Review of Political Science* 21: 403–18.

Rosenbaum, Dennis P., Amie M. Schuck, Sandra K. Costello, Darnell F. Hawkins, and Marianne K. Ring. 2005. "Attitudes toward the Police: The Effects of Direct and Vicarious Experience." *Police Quarterly* 8 (3): 343–65.

Rotberg, Robert I. 2010. *When States Fail: Causes and Consequences*. Princeton, NJ: Princeton University Press.

Rothstein, Bo. 2009. "Creating Political Legitimacy: Electoral Democracy versus Quality of Government." *American Behavioral Scientist* 53 (3): 311–30.

Ruggeri, Andrea, Han Dorussen, and Theodora-Ismene Gizelis. 2018. "On the Frontline Every Day? Subnational Deployment of United Nations Peacekeepers." *British Journal of Political Science* 48 (4): 1005–25.

Schatzberg, Michael G. 2001. *Political Legitimacy in Middle Africa: Father, Family, Food*. Bloomington: Indiana University Press.

Schmelzle, Cord, and Eric Stollenwerk. 2018. "Virtuous or Vicious Circle? Governance Effectiveness and Legitimacy in Areas of Limited Statehood." *Journal of Intervention and Statebuilding* 12 (4): 1750–2985.

Snyder, Jack L. 2000. *From Voting to Violence: Democratization and Nationalist Conflict*. New York: Norton.

Soifer, Hillel. 2008. "State Infrastructural Power: Approaches to Conceptualization and Measurement." *Studies in Comparative International Development* 43 (3–4): 231.

Stollenwerk, Eric. 2018. "Securing Legitimacy? Perceptions of Security and ISAF's Legitimacy in Northeast Afghanistan." *Journal of Intervention and Statebuilding* 12 (4): 506–26.

Sunshine, Jason, and Tom R. Tyler. 2003. "The Role of Procedural Justice and Legitimacy in Shaping Public Support for Policing." *Law & Society Review* 37 (3): 513–48.

Tyler, Tom R. 2005. "Policing in Black and White: Ethnic Group Differences in Trust and Confidence in the Police." *Police Quarterly* 8 (3): 322–42.

Tyler, Tom R. 2006. *Why People Obey the Law*. Princeton, NJ: Princeton University Press.

Viki, G. Tendayi, Michelle J. Culmer, Anja Eller, and Dominic Abrams. 2006. "Race and Willingness to Cooperate with the Police: The Roles of Quality of Contact, Attitudes towards the Behaviour and Subjective Norms." *British Journal of Social Psychology* 45 (2): 285–302.

Vinck, Patrick, Phuong Pham, and Tino Kreutzer. 2011. *Talking Peace: A Population-Based Survey on Attitudes about Security, Dispute Resolution, and Post-Conflict Reconstruction in Liberia*. Berkeley: Human Rights Center, University of California.

Weber, Max. 1962. *Basic Concepts in Sociology*. New York: Citadel Press.

Weber, Max. 1980. *Economy and Society: Ground Plan of Understanding Sociology*. Tübingen: Mohr Siebeck.

Zartman, I. William. 1995. *Collapsed States: The Disintegration and Restoration of Legitimate Authority*. Boulder: Lynne Rienner Publishers.

Ziblatt, Daniel. 2008. "Why Some Cities Provide More Public Goods than Others: A Subnational Comparison of the Provision of Public Goods in German Cities in 1912." *Studies in Comparative International Development* 43 (3–4): 273–89.