COVERT ASSIGNMENTS: UNDERCOVER INFILTRATION AND THE REPRESSION OF PROTESTS

HOWARD LIU AND BENJAMIN J. RADFORD

Abstract. What explains undercover operations in the repression of protests? The use of undercover agents to infiltrate and repress protests is highly controversial, yet their operations remain poorly understood. This research examines how the spatial characteristics of protests shape undercover agents' use of violence. We argue that undercover officers, typically lightly armed and operating in small units to avoid detection, must balance their objective to disrupt dissent against the risks of acting within dense assemblies, where arrests can provoke violent clashes and endanger officers' safety. As a result, they tend to remain passive observers and intervene only when protesters are few in number and spatially isolated, giving them a tactical advantage to arrest targets without triggering mass confrontation. Analysis of unique data from Hong Kong's largest anti-government protests (2019–20) supports these expectations: undercover officers are more likely to make arrests in locations where protesters are isolated, distant from main protest crowds or gathered in confined environments such as indoor spaces, where crowd density is low. The findings offer new insights into how covert infiltration fits within broader strategies of state coercion.

Keywords: covert repression, protest, infiltration, surveillance, undercover agents

Manuscript accepted at International Studies Quarterly in October 2025

(Howard Liu) University of South Carolina, USA
(Benjamin J. Radford) University of Pittsburgh, USA

Shai Sadeh, a prominent activist in Israel, recounted his experience of being targeted by undercover agents during a protest event. On October 10, 2020, demonstrators gathered to protest against Prime Minister Benjamin Netanyahu's corruption charges and his reluctance to resign. "I feel that there were undercover cops following me," Shai recalled. "In one of the streets they suddenly 'plucked' me, without identifying themselves, into a civilian vehicle. They filmed me with their personal smartphones and sent the pictures. I was then interrogated in the Tel Aviv station for disorderly conduct." Like Shai, other protesters shared a similar experience during the series of anti-corruption demonstrations.[1] This incident is not unique to Israel; there are numerous reports worldwide of undercover police attacking protesters and disrupting gatherings like the 2019 Hong Kong anti-government protests, inciting violence to justify crackdowns during 2021 pro-Navalny protests in Russia, and arresting activists during the 2011 Arab Spring protests in Egypt.

The deployment of undercover agents to infiltrate and repress protesters is a highly controversial practice with profound implications for human rights. Advocates for human rights decry this method, criticizing that permitting officers to masquerade as demonstrators to secretly record activities and execute arrests gravely undermines freedom of expression. Such tactics intimidate the general public, discouraging them from exercising their right due to the fear and the chilling effect it creates. And yet, the operations of undercover infiltration and repression remain poorly understood. Most existing research focuses on overt forms of repression—such as massacres (Uzonyi, 2022), population displacement (Zhukov, 2015), and public torture or beatings (Conrad and Moore, 2010)—perpetrated by visible actors like regular security forces and military. In contrast, covert forms of repression and violence performed by undercover agents remain underexplored (Ritter and Conrad, 2016; Choulis, Escribà-Folch and Mehrl, 2024). We know very little about how states use undercover agents and spies to infiltrate dissent groups and orchestrate crackdowns from within. As Davenport (2007) notes, if we continue to "ignore more covert forms of state activity, we will end up with a distorted view of what takes place."

[1] https://www.haaretz.com/israel-news/2020-10-26/ty-article/.premium/protest-leaders-undercover-cops-routinely-track-and-arrest-us/0000017f-dc00-d3ff-a7ff-fda010770000

This research presents a theoretical framework to explain covert infiltration and repression in protests, a common form of dissent. We investigate why and how undercover agents infiltrate demonstrations to perform controversial actions such as arresting protesters, often termed "sting operations" by law enforcement. We argue that badge-less undercover agents and officers serve two critical roles—*covert surveillance* and *covert strikes*—that compensate for limitations faced by their uniformed counterparts. While uniformed officers are usually armed and visible, making them efficient at deterring and repressing protesters from outside of the crowd, undercover officers operate differently. Their ability to blend into crowds due to low visibility enables them to secretly surveil and identify key agitators, and to apprehend and strike at those agitators who typically flee at the sight of uniformed police. However, their need for disguise means that undercover agents are typically lightly armed and operate in small, agile units, which puts them at a disadvantage when confronting large crowds. Intervening in dense assemblies carries heightened risks of provoking violent confrontations between protesters and police and endangering agents' own safety. To mitigate these risks while still pursuing their objectives to disrupt dissent and apprehend radicals, undercover agents generally remain passive observers and only intervene—such as by making arrests—when the group of protesters is smaller and spatially isolated, providing them with a comparative advantage in numbers. Quick "snatch-and-retreat" operations in small protest crowds reduce the likelihood of confrontation, making arrests easier and safer without escalating tensions. This logic bears important spatial implications. It suggests that undercover agents tend to make arrests in places where protesters are isolated and distant from main protest crowds or situated in spatially confined environments such as indoor spaces where crowd density is inherently low.

We test our argument using novel social media data on protests and repression in Hong Kong between 2019 and 2020. The pro-democracy protests represent the largest series of demonstrations in Hong Kong's history. They signify the most recent resistance by Hongkongers against Beijing's attempt to expand jurisdictional power and undermine the principle of "one country, two systems." Yet, demonstrators faced unprecedented repression,

leading to a significant wave of emigration prompted by fears of Beijing's retaliation against protest participants and heightened social control (Fong, 2022). Beijing deployed not just uniformed police to crack down on dissenters but also undercover officers to pinpoint activists and arrest radicalized individuals. The widespread police violence, coupled with the frequent deployment of undercover police, makes the context of Hong Kong an ideal ground to test our argument. Our data on police violence come from a popular social media platform (HKmap.live) used by Hongkongers. This platform allowed citizens to report protest events and police activity, indicating areas where protesters required assistance and where law enforcement was taking action.[2] With nearly 195,000 unique posts spanning from August 2019 to February 2020, the data offer detailed geolocation of events to the street level and timestamps by the minute. This rich dataset presents a unique opportunity to examine the spatial and temporal dynamics of undercover police operations at a fine-grained level that was not possible in prior research.

The results from our analysis support our theoretical expectations. Using panel data models and instrumental variable estimation, we find that undercover officers tend to arrest protesters as they move farther away from the main protest crowd and become increasingly isolated. We also find that undercover officers are more likely to arrest protesters in busy indoor spaces with limited room for demonstration. Both results confirm our expectations that undercover agents are more likely to engage in sting operations when they have a relative advantage in numbers and can snatch targeted individuals with minimal risk of mass confrontation.

Although the analysis of this study is confined to one country, the findings offer broader implications for research on state repression, infiltration, and secret policing in several ways. First, while the body of literature on covert repression is expanding (Sullivan and Davenport, 2018; Hager and Krakowski, 2022), existing research primarily focuses on non-violent tactics such as surveillance, leaving violent forms of covert repression like covert arrests, kidnappings,

[2]Apple, under pressure from Beijing, took down the app on October 9, 2019, stating that it endangered operations from law enforcement. However, the web version remained functional and was widely used. The Android app was also not affected.

and assassinations poorly explored. Our study demonstrates that covert operations often involve active and violent interventions to disrupt dissent, not just passive and non-violent observation by infiltrated agents. Second, we identify a distinct pathway through which governments infiltrate dissent organizations. Existing studies mainly examine secret police infiltration through recruited non-state agents, such as informants and collaborators (Scharpf and Gläßel, 2020; Nalepa and Pop-Eleches, 2022). In contrast, our study unveils a strategy of direct penetration by professional state agents who engage in both surveillance and acts of violence. Lastly, we contribute to the longstanding literature on protest-repression dynamics. While there is a growing body of work analyzing these dynamics from a spatial perspective (Olar, 2019; Christensen, 2018), most research explains repression at highly aggregated levels (e.g., countries or provinces), overlooking the local dynamics such as crowd movements in the streets and how they shape state violence. Our article takes these crowd dynamics seriously. Leveraging social media data, we investigate how the proximity of individuals to the main protest crowd and the characteristics of local gathering sites correlate with police violence.

How Governments Repress Dissent Covertly

One core question underlying many political inquiries is how we can protect human rights against state oppression. To address this question, it is crucial to understand both who states target for repression and how these targets experience systematic rights violations and abuses. This endeavor has led to rich discussions on *repression repertoires*, such as mass killing, torture, forced disappearance, and political imprisonment (Davenport and Appel, 2022; Hill and Jones, 2014; Slough and Fariss, 2021). However, since most existing work focuses on highly visible violations of physical integrity rights, we possess relatively limited knowledge about covert forms of rights violations. A recent trend, which we describe as the "invisibility turn" in the literature, pushes scholars to reconsider these covert forms of repression that, though less direct and observable, severely infringe on citizens' rights (Ritter and Conrad, 2016; Choulis, Escribà-Folch and Mehrl, 2024).

Covert repression involves subtle, indirect, and often invisible tactics of coercion, making it challenging to detect for those subjected to it. One core element of covert repression is that repressive agents intentionally conceal their identities and actions, leaving repressed individuals unaware of being targeted. Some subjects may suspect that they are being targeted but do not know who is targeting them, and may only observe agents at the moment of repression (Marx, 1988).[3] Such tactics include covert surveillance, secret harassment, and sometimes covert violence, like arrests and kidnappings. When regimes shift from overt to covert tactics, it typically indicates that regimes are taking a more proactive approach to seeking out and neutralizing dissent rather than merely waiting and reacting to challenges (De Jaegher and Hoyer, 2019; Liu and Sullivan, 2021). This shift often requires infiltration into dissent groups to gather critical information about dissent activities.

Covert repression can be categorized into nonviolent and violent forms, with more work focusing on the former. The most common type of nonviolent covert operation is *covert surveillance*, where agents infiltrate protest movements and dissent organizations to collect information on past, planned, or ongoing dissent activities (Steinert, 2023; Thomson, 2023).[4] Agents in these operations are generally passive, engaging primarily in observation, monitoring, and questioning. They act as an "informational sponge" and like "the fly on the wall," aiming to gather information without altering the behavior of their targets. While much research focuses on surveillance conducted by non-state agents—such as civilian informants,

[3]This study focuses on the covertness of agents, but we acknowledge that there are other dimensions of covertness that can be situated in the concept of covert repression. One such dimension is government acknowledgment: political disappearances are often considered covert because the state denies detaining or killing the victims, leaving families uncertain about victims' fate. Future research would benefit from explicitly identifying which dimension of covertness is under examination, as doing so can clarify distinctions among different forms of covert repression and help theory-building.

[4]Another commonly used nonviolent covert tactic is decomposition measures (Zersetzungsmaßnahmen)—a form of covert harassment employed by the Stasi's undercover agents in the former GDR (Steinert, 2023; Steinert and Dworschak, 2023). These agents can secretly create conflicts between spouses by sending anonymous 'love' letters or creating division among dissident group members by spreading rumors.

collaborators, and regime sympathizers (Piotrowska, 2020; Nalepa and Pop-Eleches, 2022)– professional state agents, including intelligence officers and specialized police force,[5] also play significant roles in spying on targets (Scharpf and Gläßel, 2020).

Existing research has paid much less attention to violent forms of covert repression, which we term as *covert strikes*. In these operations, agents infiltrate to identify and neutralize dissidents, often engaging in active intervention to alter the behavior of their targets so as to prevent, limit, and stop dissenting activity (Hassan, Mattingly and Nugent, 2022). These tactics include covert arrests, kidnappings, and sometimes even extreme violence like secret assassinations. Infiltration is often employed to get closer to hard-to-capture targets unlikely to be apprehended through conventional means. Due to the violent and sometimes illegal nature of these acts, agents of repression often carry out these tactics discreetly to avoid public attention and accountability. Examples include the FBI's COINTELPRO operation during the 1960s, which targeted black nationalist movements with covert arrests, kidnappings, and secret assassinations, and Operation Condor in the 1970s, where South American states covertly kidnapped and assassinated key anti-regime figures (Cunningham, 2003; Starr et al., 2008). Because covert violence requires high-level skills, professional state agents are more commonly involved than untrained civilians and informants. Although some existing research has employed qualitative approaches such as descriptive case studies and interviews to provide early insights into covert operations, data on these operations are difficult to collect, and quantitative research on this topic remains limited.

In this article, we explore the under-researched but important tactic of covert strikes, examining why infiltrated agents sometimes covertly strike targets while at other times they remain passive observers. By focusing on the security force's infiltration of protest events, we show that the dynamics of protest congregations on the ground correlate with undercover agents' strategies between strikes and surveillance. Studying these tactics advances our

[5]While some research tends to separate secret police from the regular police in discussing covert repression, we agree with Marx (1988) in viewing little difference between these two units when agents target political groups.

understanding of how governments incorporate covert infiltration into their strategies of coercion.

The Logic of Undercover Infiltration, Violence, and Repression of Protests

In discussing the strategic behavior of covert agents, it is essential to differentiate between two types of covert operations—deep and light infiltration—due to their associated risks. In *deep cover operations*, undercover agents immerse themselves within target organizations for extended periods, often adopting entirely new identities. This comprehensive immersion requires a social rebirth, where agents live a new life under a constructed persona. Such a high level of integration demands rigorous and comprehensive training, which includes developing a believable backstory, mastering encrypted communication to evade detection, and acquiring advanced investigative skills (Fijnaut and Marx, 1995). The primary aim of deep cover operations is to forge intimate interpersonal relationships, gaining trust to access critical insider information about organizational plans, logistics, and leadership—details typically inaccessible through external surveillance alone (Liu, 2022). These operations allow agents to subtly influence the organization's direction or coordinate with other security force units to disrupt dissent activities. However, deep cover operations carry substantial risks; a blown cover can have fatal consequences, given the depth of integration and the long-term commitment to the covert identity. Therefore, agents in deep cover operations usually focus on covert surveillance, avoiding aggressive actions to mitigate the risk of exposure and protect their safety.

Light Cover Operations, in contrast, are designed for short-term immersion to address immediate threats, often within the context of protest policing or quick disruption of smaller-scale dissident activities. Infiltrated agents do not live a new life and their cover is usually for a temporary assimilation. These operations require agents to blend in temporarily, aiming for rapid identification of key dissidents and the interruption of specific, smaller-scale illegal or dissenting acts. Infiltration is particularly useful for apprehending hard-to-arrest radicals

who often blend into crowds and flee at the first sign of police presence. The risk of identity exposure in these operations is generally lower due to agents' less integrated roles, a quick exit strategy in mind, and possible backup support compared to deep cover operations. Therefore, agents in light infiltration engage in more direct interventions, such as covert strikes, against targeted activities, allowing for quick and effective responses to emerging threats. The following quote from one of the Miami Police's protest infiltration reports vividly captures the key roles undercover agents play in protest management:

> *"Plainclothes and the so-called 'extraction teams' were created for assignment inside the demonstrations or marches. For the most part, these plainclothes officers were to assume only an observation role. However, these officers were also trained when to take police action, which included situations where a radical group member was causing significant property damage, committing a felonious act, or attacking innocent people...The extraction teams were intended to separate violent actors from within crowds without escalating a larger confrontation between the crowd and police. This dangerous assignment required swift action by the plainclothes officers in the face of expected attacks by those seeking to free their compatriots."*
>
> — *City of Miami Police Department, FTAA: After action review, 2004.*

However, it does not mean that undercover agents will always engage and take intervening actions. Aggressive measures against protesters amidst a large gathering remain risky for these agents. Due to their need to disguise, undercover agents are often lightly armed and operate in small, mobile teams to avoid detection. This can put them at a disadvantage and potentially in danger, given their reduced numbers, especially when apprehending individuals within dense assemblies. Protesters may become enraged when discovering that state agents, such as undercover police, are covertly embedded within their ranks and spying on them, incentivizing protesters to disrupt apprehension or even attack isolated undercover agents to free their compatriots, as mentioned in the Miami police report. To mitigate these risks, undercover agents need to maintain their roles as observers in large crowds and defer

intervention to uniformed officers even though it means that they cannot immediately disrupt radical behaviors and risk losing track of radicals when agents wait for backup reinforcements.

Intervening actions become safer and more feasible when the number of protesters decreases. Smaller crowds provide undercover agents with a tactical advantage, reducing the risk of resistance and the likelihood of triggering violent clashes between the crowd and police during arrests. Undercover officers can more easily execute quick snatch-and-retreat tactics that disrupt radical behaviors and arrest radical individuals while minimizing confrontation. As the number of protesters drops, perceived risks to undercover agents decline, thus making arrests more likely. This logic has two spatial implications: arrests executed by undercover agents tend to occur (1) in locations where individual protesters are spatially isolated and distant from the main gathering sites, or (2) in confined environments (e.g., indoor spaces) where crowd density is inherently low. We will further elaborate on this strategic logic in the following sections.

Distance from the Main Protest Crowd. Physical control of space is fundamental to protest policing and crowd control (Kaufman, 2016; Bass, 2001). The location of protesters directly determines where security forces intend to go and the spaces where authorities intend to control. That is why we often see repression and arrests happen within the main crowd of protesters, especially when protesters remain in a space after being ordered to vacate. Existing studies often refer to these gathering sites as "hotspots" and employ spatial hotspot analysis to predict locations where police may resort to force (Braga, Papachristos and Hureau, 2014; Weisburd and Braga, 2006). Confronted with a large group of defiant protesters, police often employ physical force to disperse the crowd. In this context, both uniformed and undercover officers play pivotal roles in crowd dispersal. Uniformed officers equipped with tactical gear (e.g., helmets, body armor vests, shields, and batons) are better prepared for both offensive and defensive actions against protesters. But gear compromises mobility. Undercover agents like plainclothes officers do not wear heavy gear and are more mobile; this facilitates chasing and arresting individual protesters. When plainclothes police take action,

operating near uniformed police can offer protection and increase their safety. Collaboration between uniformed and undercover police makes mass arrests more effective and crowd control more efficient.

While mass arrests can quickly control and disperse crowds, this strategy is often messy, highly confrontational, and violent. Moreover, indiscriminate repression frequently leads to the accidental apprehension of innocent bystanders, resulting in backlash and increased resistance (McPhail and McCarthy, 2005; Lyall, 2009).[6] To address this issue, law enforcement sometimes adopts a more targeted approach to reduce mass confrontation and violence at protest sites: undercover officers identify and discreetly follow radical protesters, snatching them when they are alone and away from the main protest crowd.

This strategy offers several tactical advantages. First, when individual protesters step away from the main demonstration sites—whether to rest, eat, or head home—they become increasingly isolated and more vulnerable to apprehension.[7] The farther they are from these main zones of gathering, the fewer fellow protesters are nearby, making them easier and safer targets for arrest. Undercover agents, typically operating in small roving teams to avoid detection, can easily take down a few unarmed protesters when they have the numbers advantage. Second, making arrests away from the main protest crowd limits the potential for collective resistance. With fellow protesters too far to intervene, undercover agents face a lower likelihood of arrest interference, disruption, or attempts to rescue the target. This spatial separation also reduces the chance of escalation into broader confrontations between protesters and police during arrests. Moreover, isolated individuals face greater difficulty escaping back to the crowd or seeking refuge, further increasing their vulnerability. Third, arrests conducted away from high-visibility protest frontlines or hotspots reduce the likelihood of police violence being recorded and broadcast by the media. Under light cover

[6]The backlash effect is a well-recognized phenomenon and has been widely examined in the context of repression and protest. See, for example, Steinert and Dworschak (2023), Schulte and Steinert (2023), and Demirel-Pegg and Rasler (2021).

[7]When we say protesters are isolated and alone, we do not mean that there are absolutely no people nearby. We simply mean that there are few or no *fellow protesters* in their vicinity who could support the targeted protester or intervene to obstruct an arrest, as these targeted protesters are farther from the main protest crowd.

missions such as crowd control, undercover officers are less afraid of exposure and, in fact, often identify themselves during the arrests to assert legal authority and deter obstruction. However, they prefer to avoid public scenes when their actions involve aggressive violence and cause injury to arrested individuals. By targeting individual protesters in less monitored areas rather than the frontline protest zones, law enforcement can preserve a more restrained image when performing crowd control. Therefore, we derive the following hypothesis:

> *H1: Undercover agents are more likely to arrest protesters as their distance from the protest crowd increases.*

Gatherings in Indoor Spaces. When we think of protest, we typically associate it with demonstrations in public, open outdoor spaces such as streets, parks, or city squares. But when governments are repressive and gatherings in public outdoor spaces are restricted, protesters are forced to be creative and select other places where they can still convey their message to large audiences. Semi-public indoor locations where people assemble (e.g., airports, subway stations, sports stadiums, and shopping malls) have become new venues for protesters in recent years (McCarthy and McPhail, 2006). Although these confined environments limit the number of protest participants, they allow protesters to broadcast their message by engaging with large flows of passersby. Such behavior can help mobilize public support and raise awareness. For example, in 2017, protesters gathered at airports in the USA, including JFK in New York and O'Hare in Chicago, to voice opposition to the travel ban that targets seven predominantly Muslim countries. During the 2019 Cricket World Cup in England, Afghan fans protested against the Pakistani government's policies in relation to Afghanistan in football stadiums. In 2014, protesters gathered at the Mall of America in Bloomington, Minnesota, to protest the killings of unarmed black men by police. These examples illustrate how semi-public indoor venues have become critical arenas for dissent when traditional protest spaces become restricted or less accessible.

At the same time, protests in confined indoor spaces create distinct challenges for law enforcement. Gatherings in busy indoor areas make it difficult for police to distinguish between radical individuals, peaceful protesters, and uninvolved bystanders. This difficulty increases when radicals engage in hit-and-run tactics to vandalize property or commit violence. Responding with mass, indiscriminate arrests can harm bystanders and passersby, cause public fear, and disrupt businesses and daily activities that governments often seek to avoid. A more targeted approach focused on identifying and neutralizing radical individuals is preferable. Undercover infiltration and covert strikes are particularly well-suited to this type of operation.

These operations benefit from several strategic advantages in confined indoor environments. First, indoor spaces naturally limit the size and density of protest gatherings. Compared to protests in open outdoor spaces, these smaller and more dispersed crowds give undercover agents a tactical advantage. They offer undercover agents an upper hand to apprehend radicalized individuals with less resistance. For example, locations like subway stations and shopping malls provide limited space for assembly and are often filled with uninvolved bystanders and passersby. Protesters in these busy venues tend to be small in size and often scattered around the space, making them more vulnerable to targeted arrests. Second, arrests in such settings are less likely to trigger large-scale confrontations. Because these spaces are physically separated from outdoor protest gatherings, the risk of arrest interference from other demonstrators is reduced. Moreover, although these indoor venues may be crowded, the surrounding individuals—commuters, shoppers, or pedestrians—are not part of the protest and therefore unlikely to intervene or obstruct police action. This combination of physical separation and uninvolved bystanders lowers the likelihood of arrest obstruction and makes arrests safer and more manageable for undercover agents. Lastly, semi-enclosed indoor spaces often have limited entry and exit points, enabling undercover officers to ambush protesters at chokepoints such as stairwells, escalators, or gateways. These physical constraints in space make it harder for protesters to escape, giving officers greater control over the arrest environment.

It is worth noting that while undercover officers generally prefer to avoid public attention—particularly to minimize documentation of their use of violence that can tarnish the public perception of police—they do not necessarily seek to remain entirely unseen throughout all phases of their operations-covert surveillance, covert strike (arrest), and retreat. In light cover operations such as short-term protest infiltration, undercover agents often *identify themselves at the moment of arrest* to establish legal authority and deter obstruction from nearby citizens or journalists. Their primary concern is not avoiding exposure when they make arrests, but rather ensuring tactical advantage and their own safety during the arrests. If they can carry out easy and safe arrests — even in somewhat visible locations like malls or subway stations — they will often accept some degree of public exposure and take intervening action. Ultimately, what matters most for officers is not avoiding attention at all costs, but securing a manageable arrest environment and preventing conflict escalation.

Based on the strategic advantages that semi-public indoor venues offer to undercover agents—such as smaller protest crowds, reduced risk of interference, and greater physical control over space—we derive the following hypothesis:

> *H2: Undercover agents are more likely to arrest protesters in indoor spaces that experience a high flow of people and passersby (e.g., airports, subway stations, and shopping malls) as opposed to other locations.*

Empirical Context: Anti-government Protests and Repression in Hong Kong (2019-20)

In 2019, Hong Kong experienced a major outbreak of anti-government protests against the Extradition Law Amendment Bill. Also called the anti-ELAB protest, it was triggered by a proposed bill that would have required Hong Kong to send fugitives to Mainland China without existing bilateral extradition agreements with Hong Kong, including mainland China, Taiwan, and Macau. The bill proposed by Hong Kong's pro-Beijing administration would reinforce jurisdictional power for Beijing to persecute political dissidents, which undermined Beijing's promise of a specialized governing system "one country, two systems" for Hong

Kong. Despite widespread public concerns, the bill was set to pass the second reading in Hong Kong's Legislative Council (LegCo) on June 12, 2019. In response, tens of thousands of protesters surrounded the LegCo and forced the cancellation of the meeting. While the government suspended the bill on June 15, the protest escalated to incorporate other demands including the complete withdrawal of the bill and the retraction of the riot characterization of the June 12 protest. The protests were met by brutal police violence and eventually turned into weekly pro-democracy protests until February 2020 when the COVID-19 pandemic and city-wide lockdowns brought them to a recess.

The anti-ELAB protests were often described as leaderless movements in which protesters did not follow specific organizers or leaders but chose instead to join protests spontaneously based on online mobilization calls. The decentralized nature of the protests stemmed from lessons learned by protesters during the government's targeted repression against the leaders of Hong Kong's Umbrella Movement in 2014. Without leaders, the anti-ELAB protesters coordinated their activities through messaging apps likes Telegram and online discussion forums, such as LIHKG (a Reddit-like discussion forum). Movement activists also created a website called *HKmap.live*. It is a crowdsourced online mapping service and app where information about police movements and strategies were posted to alert fellow protesters of police presence and to help protesters avoid danger zones. This information came from civilian and protester reports which were verified by the website administrators before being posted online.

The extent of police violence and brutality was unprecedented in the history of Hong Kong after its return to China in 1997. The lack of a political solution between protesters and the government undermined the peaceful prospects and led to the forceful dispersal of crowds. The Hong Kong government, under pressure from Beijing to preserve order, also fueled the frequent and severe use of police violence. The international community openly condemned the human rights violations witnessed during this period, including suppression of freedoms of assembly and expression, arbitrary arrest and detention, and excessive use of force against

unarmed protesters.[8] Riot police, fully armed and in green suits/uniforms, became a common sight. The police also relied heavily on plainclothes police to surveil protesters and, more importantly, to initiate arrests of protesters when they were unguarded. The pervasive use of undercover agents in managing primarily unarmed demonstrators emerged as a major concern within Hong Kong society.[9] The widespread police violence, coupled with the frequent deployment of undercover police, makes the context of Hong Kong an ideal ground to test our argument on undercover policing and protest repression.

Research Design

Data. We collect policing data by scraping police and protesters information from one of the most popular social media sites (HKmap.live) used by Hong Kongese during the protests. Figure 1 shows a snapshot of data reported on this site. Although the protest movement started in June 2019 and the site was not launched until August 4th 2019, the data we collect cover a majority of incidents during the most contentious times all the way until the end of the movement in January 2020 before the COVID lockdowns. During this time, we ran our scraper to collect the most up-to-date reports and removed duplicates using the provided report IDs. Because we ran our scraper in real-time, we were able to analyze all historical data despite the site not providing access to them.

Scraping a dynamic web-based mapping service like this is not an easy task. The site deployed *Cloudflare*, a service designed to block content scraping and mitigate distributed denial of service (DDoS) attacks that would otherwise interrupt access to the site for legitimate users. This firewall created challenges for our scraping.[10] We took two steps to overcome this: (1) we modified our scraper to collect updates every three minutes (less likely to trigger countermeasures) rather than every minute, and (2) we rewrote our scraper to use

[8] https://www.cnn.com/2019/11/19/asia/hong-kong-senate-democracy-bill-intl-hnk/index.html

[9] https://www.theguardian.com/world/2019/aug/12/hong-kong-protests-brutal-undercover-police-tactics-spark-outcry

[10] After several days of scraping, our initial web scraper triggered Cloudflare's countermeasures which implemented a browser check for requests originating from our IP address (and others) and effectively banned our web scraping software.

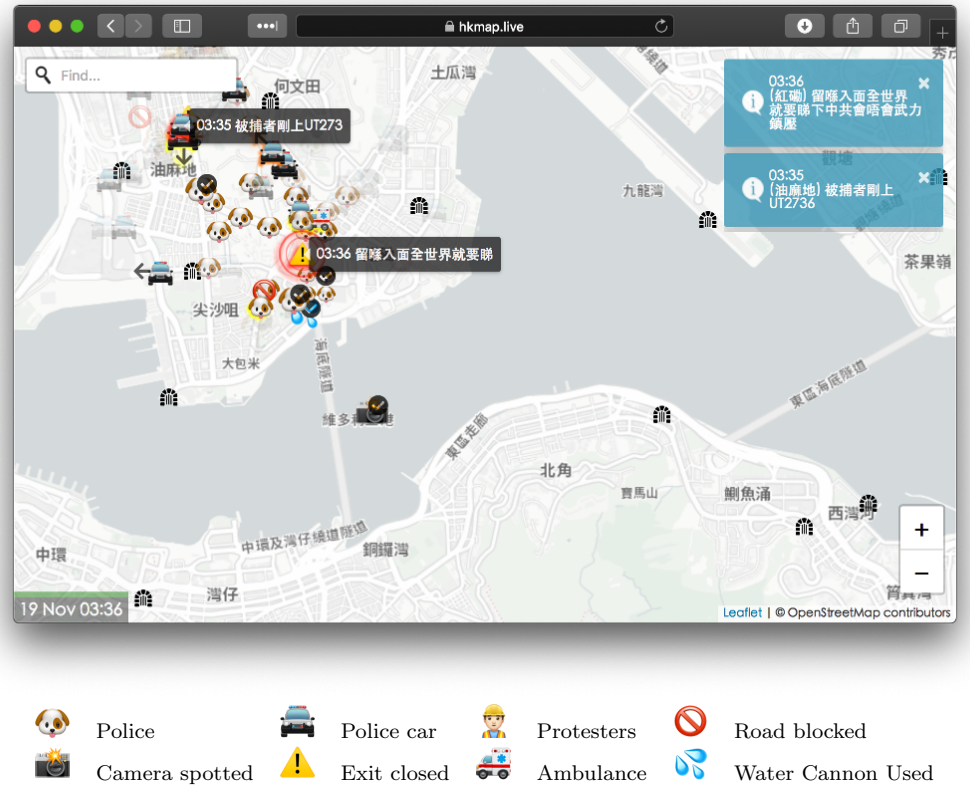

Figure 1. A snapshot of visualized data from HKmap.live. The legend provides a general interpretation of these icons on the map, although their specific meaning may be context-dependent.

a headless browser through Selenium, an API for automating browser actions. The headless browser allowed our scraper to wait until after the browser check had been completed and then rerouted our request to HKMap.live content. With this design, our scraper was able to access posted content without disruption.

It is worth noting that initially, our delayed discovery of this mapping site and setup of a sophisticated scraper led to some missing data for us from much of August and early September. However, we obtained these missing data directly from one of the site administrators who graciously shared the missing portion of the data we need.[11] Additionally, the several-day overlap between the administrator-provided data and our scraped data offered a unique opportunity to validate the completeness of our collection efforts. The result shows that we successfully captured at least 99.1% of all posts and at least 99.4% when accounting

[11]He shared the data with us after he had left Hong Kong safely.

for data left-censored in our records. It strengthens our confidence in the comprehensiveness of the scraped data.

Data from social media provide several advantages over traditional sources. Previous studies on state repression predominantly utilized event datasets generated from news reports (Bagozzi, Berliner and Welch, 2021; Hendrix and Salehyan, 2017) or specialized data generated from declassified government archives (Sullivan, 2016; Liu, Su and Wang, 2025). This is because governments seldom disclose information on their repressive actions. While declassified documents serve as valuable governmental records, they are usually released long after the actual events or after a regime change. Thus, they are more suitable for studying historical rather than contemporary events. News reports, on the other hand, have been widely used to study state repression. But they are known to have a bias toward larger and more violent events as well as events in urban areas with greater information accessibility (Weidmann, 2016).

While social media data can still suffer from urban-rural biases given varying internet penetration rates, this problem is mitigated in the context of Hong Kong. It is an international city with one of the highest population concentrations in the world. Internet penetration is high and smartphones are prevalent. With the help of self-organized civil groups collecting and reporting pro-democracy protests and policing activity on social media, we can observe data that provide more details on timing, location, and micro-level actor identities. Figure 2 depicts the spatial distribution of events in Hong Kong during the period from August to December 2019. The inset illustrates posts within a quarter-kilometer square area throughout the 12-day siege of Hong Kong Polytechnic University in November. A subset of the data released by HKMap.live administrators have been used to study the police's use of tear gas in published research (Chau and Wan, 2022), bolstering our confidence in this data source. We also cross-referenced this data with news reports. The results showed a strong correlation, but HKmap.live had the advantage of capturing a greater number of events, spanning diverse locations and smaller scales. We detail our validation in the robustness section.

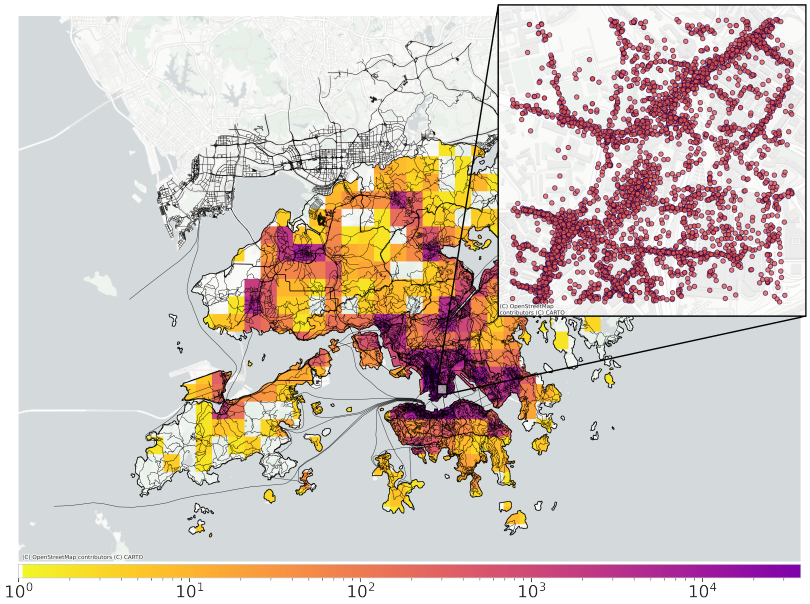

Figure 2. The distribution of protest events between August and December, 2019 (outer) and the siege of Hong Kong Polytechnic University (inset).

Unit of Analysis. Our study investigates the link between spatial distances, location features, and undercover police actions. Leveraging the spatially fine-grained observations from the social media app, we compile data at the level of district council constituencies–Hong Kong's smallest administrative divisions, comprising 452 constituencies in a densely populated city. Temporally, given the dynamic nature of protest crowds and marches, which can change composition significantly throughout the day, it becomes challenging to measure individuals' distances from these crowds on a daily level. To address this, we compile data at the more refined quarter-day level to better track the dynamics in protest crowd formation and the spatial distances of individual protesters from these crowds within a day.

Measurements. Our dependent variable is a binary indicator that denotes whether undercover officers make arrests.[12] We source our data from HKmap.live, which comprises short comments describing the observed events, and we rely on the comments to code events related to undercover officers' arrests.[13] It is coded as 1 if there were any arrest events made by

[12]Binary measures are used to address the potential concern about duplicated arrest reports that use slightly different times and geolocations.

[13]HKmap.live includes brief comments with reported events, and we use these comments to discern who carried out the arrests. Sample comments include: "Comrades, be cautious. Many plainclothes officers in grey shirts are arresting people!" (手足, 重量注意, 大量灰衫便衣捉人！), "Two comrades were caught by

undercover police or events where the presence of undercover police and arrests co-occurs in the same electoral district and the same quarter of the day.[14] Otherwise, it is coded as 0. It is worth noting that while plainclothes police are hard to observe given their disguise, they can be easily identified when they take police actions, especially when arresting protesters. In the case of Hong Kong, there is no evidence of vigilante groups disguising themselves as cops to arrest protesters like in the US, so it is reasonable to assume that the actions observed were conducted by real undercover police officers.

We have two explanatory variables for each hypothesis. To test the first hypothesis, we measure protesters' relative distance from the main protest crowds. This distance measure serves as a proxy for the number of protesters in a given location, as greater distance from the main protest crowd implies fewer nearby fellow protesters—indicating a smaller protest presence and greater spatial isolation of individuals from the main site of gatherings. Creating this measure requires two steps: first, identifying the main protest crowds, and second, calculating protesters' spatial distance from these crowds. This process poses two empirical challenges: identifying the "main crowd" of protesters spatially on the ground, and ensuring that the measure is dynamic over time. To address these challenges, we apply an unsupervised machine learning method to identify clusters of protesters both spatially and temporally.

Specifically, to identify the main protest crowds (which we empirically refer to as *main protest zones*), we apply a density-based clustering algorithm, DBSCAN (Density-Based Spatial Clustering of Applications with Noise). DBSCAN detects clusters of spatial points (i.e., reported protest events) based on their spatial concentration. It is a widely used method in spatial data analysis that partitions large datasets into clusters of high point density, thereby identifying dense regions in the data space. These clusters represent areas with the highest density of protest activity in our analysis. By definition, the farther a protester is away from these protest clusters or zones, the lower the surrounding density of

plainclothes officers." (2手足被便衣咬), and "Three plainclothes officers are making arrests inside the mall." (3便衣商場入面逗人).

[14]Co-occurrence is considered as there are some instances where reports did not mention who made the arrest.

protest events. To capture clustering dynamically, we set the ϵ parameter—the maximum possible distance between two data points—to one kilometer in space and one hour in time. This ensures that any point in a cluster is within these thresholds from its nearest neighbor. DBSCAN is well-suited for this task because it identifies clusters of varying sizes and irregular shapes while being robust to noise and outliers (Bin Waheed, Al-Zahrani and Hanafy, 2019). Unlike other clustering methods such as K-means, DBSCAN does not require pre-specifying the number of clusters, allowing the clustering structure to emerge from the data itself. An example of a protest cluster identified by this method is shown in Figure 3.

Figure 3. One example of protest zone clusters identified in the afternoon of October 20, 2019 in the Kowloon area. Colored circles denote reports on HKMap.live that directly indicate the presence of protesters. The main protest zone cluster is highlighted by an enclosing polygon that is provided for illustration purposes.

After the main protest zone clusters are measured, we then proceed to calculate the relative distance to the main protest zones, *Closeness to protest zone*, by the following formula:

$$Wz_t = \sum_{j}^{n} W_{i,j} \times z_{j,t-1}$$

where $w_{i,j}$ is an inversely weighted spatial distance matrix measuring the distance between each pair of spatial units i and j. Inversely weighted distance matrices ensure that closeness decays as the distance grows. w_{ij} is then multiplied by the zone indicator variable $z_{j,t-1}$ which denotes whether a spatial unit j is located within a zone. The zone indicator, like common spatial-temporal lag variables, is lagged by a quarter day $t-1$ to avoid simultaneity

concerns. This weighted distance provides an aggregated measure of how close a spatial unit is to the main protest zone at the quarter-day level.[15]

To test our second hypothesis, we need to identify indoor spaces where protesters can gather and also experience a relatively high flow of people and foot traffic. This excludes locations inaccessible to protesters, such as private business buildings and residential communities that are strictly private and require identification for entry, as well as public indoor spaces like courthouses, museums, libraries, and hospitals, where the public can access but protests and assemblies are strictly restricted. According to research, popular indoor locations that attract crowds and where protesters frequently assemble in Hong Kong during the wave of protests include underground subway stations (also known as the Mass Transit Railway, MTR), the Hong Kong airport, and indoor shopping malls (Stokols, 2023). We collect geolocations for these sites from the Open Street Map Database and Wikipedia, and identify the electoral constituency units where these spaces are situated, coding them as 1 if such a space is present and 0 otherwise.

Additionally, we recognize that comparing spatial units to investigate the relationship between indoor locations and arrests by undercover police can lead to an issue where arrests that actually occur outside these indoor spaces are considered indoor arrests, which will undermine our test on the logic of indoor-oriented undercover arrests. To address this, we also utilize the centers of these locations to create spatial bubbles, approximating the boundaries of these locations in our test. These analyses will be further elaborated in the robustness section.

Model. Of course, protesters' distances from the main protest crowds are not randomly assigned across our spatial units and there is reason for concern that whatever relationships emerge are confounded by other factors. There are two broad concerns here: one is that other factors –such as observed police presence and the incidence of violence–affect both protesters' relative distances to crowds and undercover arrests. The other concern is reverse

[15]While it is empirically more robust to measure main protest zones as it directly speaks to our theory, our result remains consistent when proximity is measured by calculating the average distance to all other protest events (without a zone identified) for each spatial unit.

causality: the incidence of undercover arrests affects where protesters choose to go, not the other way around. We attempt to overcome these challenges in several ways, though we note here that these challenges are fundamental to observational work. Our results should primarily be interpreted as correlational rather than causal.

The first way we address confounding is through controlling for observables. First, since we are interested in exploring how proximity to the main protest crowd affects arrests, we include the location of the *Protest zone* as controls. Second, we include a variable, *Arrest nearby*, to account for the potential diffusion effect in police arrest. Third, where police, either uniformed or undercover, are located may also affect the likelihood of arrests, so we include two indicator variables to measure whether *Undercover police* or *Uniformed police* are observed. Additionally, we include a variable, *Protest violence*, to account for the possibility that police may be more willing to repress when protesters are engaged in violence.[16] We also include fixed effects to account for unit-specific characteristics that may also affect the behavior in undercover policing. Finally, for the indoor spaces hypothesis, we also include the location of the Legislative Council, *LegCo*, which is the location of the Hong Kong Special Government to account for the fact that arrests may be more likely in the political/administration center to ensure security. We also account for temporal dependencies in our observations by incorporating time polynomials (Carter and Signorino, 2010).

Our second approach to address confounding is to employ the instrumental variable design. Because the distance to main crowds of protest is affected by crowd size, and crowd size can be endogenous to undercover policing activity, a possible reversed causal argument exists: undercover police arrests may reversely influence crowd size. The reports of undercover surveillance and arrests could deter people from attending protests due to fear, or conversely, reports of undercover officers attacking and arresting protesters could incite anger over the issue of injustices and excessive use of force, prompting more people to participate in protests.

[16]HKmap.live data do not have information on protest violence. We manually coded instances of protest violence by reading through a newspaper archive, collected by (Teo and Fu, 2021), that documents protest events and event descriptions from August 2019 to January 2020.

To address this reverse causality concern, we apply a two-stage least squares (2SLS) instrumental variable approach. An instrument should predict the main independent variable of interest (crowd size, in our case) but should not be causally related to the dependent variable (arrests by undercover officers) except through its effect on the independent variable. We follow the recent innovation in the literature that utilizes time, specifically *the day of the week*, as an instrument. Butcher and Pinckney (2022) use Fridays as an instrument to study the relationship between protest size and government responses in Muslim countries. We extend this approach in our study of Hong Kong where citizens are more likely to protest during weekends when they do not have to work or go to school.[17] Because our data present substantial variation of crowd size within a day, we also include *the time of the day* as our second instrument. Citizens are more likely to protest during the daytime rather than late at night.

For the exclusion restriction assumption to hold, time must directly predict protest mobilization but not arrests. We argue that this assumption is reasonably held in our study. Weekends increase the likelihood of protest crowd size, but they are unlikely to directly increase arrests, except through increased crowd sizes. There are no laws or regulations banning protests from happening during certain days of the week, making it unlikely that police would make more or fewer arrests on weekends irrespective of protest crowd size. A similar logic applies to the time of the day. Protesters are more likely to gather during the daytime when they are awake, but days or nights should not directly increase or decrease police arrests regardless of protest size.

However, concerns may arise that police officers may feel upset because they have to work on weekends or at night and thus act more aggressively and become more likely to engage in arrests due to their anger. In other words, mood or emotion could serve as a causal pathway between time and police arrest decisions. We could not find any data measuring the mood of police officers in Hong Kong during our study period. We did, however, obtain an official daily arrest record (from 06/2019 to 01/2020) requested by a pro-democracy lawmaker, Eddie

[17]We also include Fridays because they are considered a part of the weekend when Hong Kongese often finish work or school in the afternoon, providing an opportunity to join protest.

Chu Hoi-dick, and later released by the Hong Kong Security Bureau in February 2020. We analyzed the frequency of arrests by days of the week and presented a bar chart in Appendix Figure A.1. We see no evidence of increased arrests on weekends. Arrest days are quite evenly distributed throughout the week. The concern that police officers tend to engage in arrests on weekends due to anger is thus less likely. Additionally, concerns may be that arrests might increase at night when protests are likely to be more violent. To address this, we collected data on protest violence from news reports. Our findings remain consistent after accounting for protest violence.

In testing our second hypothesis, the reversed causality concern is less relevant since our measurement of indoor spaces is both temporally and spatially consistent throughout the duration of our analysis. We employ a random-effect model, rather than fixed-effects, given the time-invariant nature of the indoor spaces variable.

Results

We estimate several models to test our two hypotheses. The results are presented in Table 1. We start from columns one and two. In both least squares models, one with a time polynomial term (2) and the other without (1), we find support for Hypothesis 1. In particular, the spatial term, *Closeness to protest zone*, obtains a negative coefficient. In other words, controlling for the arrests that occur within the main *protest zone*, we find that undercover police become more likely to arrest protesters as protesters' distance from the main protest crowd increases. The result remains robust when we include time polynomials to account for temporal dependencies among observations.

However as previously mentioned, protest zones may be endogenous to undercover police activity. Therefore, we estimate two additional models, shown in columns three and four, that instrument for protest zone, *Predicted protest zone*, and closeness to protest zone, *Predicted closeness to protest zone*. These columns are the second stages of 2SLS models, and the first stages are reported in Appendix Table A.1. We find that the 2SLS approach concurs with the initial OLS models. Even when instrumenting for protest zones and closeness to protest

	OLS	OLS	2SLS	2SLS	OLS	OLS
Closeness to protest zone	−0.0640***	−0.0649***			−0.0649***	−0.0658***
	(0.0130)	(0.0131)			(0.0029)	(0.0029)
Protest zone	0.0571***	0.0571***			0.0576***	0.0576***
	(0.0068)	(0.0068)			(0.0007)	(0.0007)
Undercover Police	0.1818***	0.1818***	0.1378***	0.1376***	0.1824***	0.1824***
	(0.0115)	(0.0115)	(0.0183)	(0.0183)	(0.0007)	(0.0007)
Uniformed Police	−0.0029***	−0.0029***	−0.0100***	−0.0100***	−0.0027***	−0.0027***
	(0.0005)	(0.0005)	(0.0025)	(0.0025)	(0.0002)	(0.0002)
Arrests nearby	0.0523***	0.0525***	0.0439***	0.0441***	0.0520***	0.0522***
	(0.0113)	(0.0113)	(0.0121)	(0.0120)	(0.0015)	(0.0015)
Protest violence	0.0128	0.0129	0.0038	0.0038	0.0124***	0.0125***
	(0.0123)	(0.0123)	(0.0073)	(0.0073)	(0.0012)	(0.0012)
Predicted closeness to protest zone			−0.2456***	−0.2473***		
			(0.0612)	(0.0613)		
Predicted protest zone			0.3407***	0.3424***		
			(0.0902)	(0.0907)		
Indoor spaces					0.0004**	0.0004**
					(0.0002)	(0.0002)
LegCo					0.0039**	0.0038**
					(0.0016)	(0.0016)
Model	FE	FE	FE	FE	RE	RE
TimePoly	N	Y	N	Y	N	Y
Num. obs.	327248	327248	327248	327248	327248	327248
Num. groups: CACODE	452	452	452	452	452	452

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 1. Effects on Undercover Policing

zones, we find that protest zones are associated with increased undercover arrests while the spatial term indicates a negative relationship between undercover arrests and closeness to the protest zone, supporting our Hypothesis 1. All first-stage instruments achieve F-statistics of at least 43 to and up to 1,936, which are substantially larger than the conventional threshold of 10 and thus show no evidence of weak instruments (Sovey and Green, 2011). Turning to models 5 and 6, our findings show that undercover police are more likely to make arrests in indoor spaces, supporting our Hypothesis 2. Our results remain consistent when applying a logit model with random effects, as shown in Appendix table A.2.

Robustness Checks

One may be concerned that our data from HKmap.live are not representative because they are reported by citizens. To validate our data, we compare our event data with news-reported event data. We hand-coded news-reported protest events from Wikipedia's "Timeline of the

2019-2020 Hong Kong Protests"[18] and included events that were not mentioned on Wikipedia but included in a comprehensive collection of protest news reports from a research team at the University of Hong Kong.[19] The distribution of protests from both sources over time is given in Figure 4. The comparison shows high similarity in the trend, bolstering our confidence in our data. It also shows that HKmap.live reports more events than the news. This is because the news tends to only report the largest demonstrations per day, whereas HKmap.live reports more detailed information for events that are smaller in scale and occur in multiple locations within a day, which are essential for our analysis. This comparison indicates that HKmap.live not only aligns with the general trends of conventional news but also offers more granular, daily, and multi-location event data, presenting a more comprehensive view of Hong Kong's protest and policing activities.

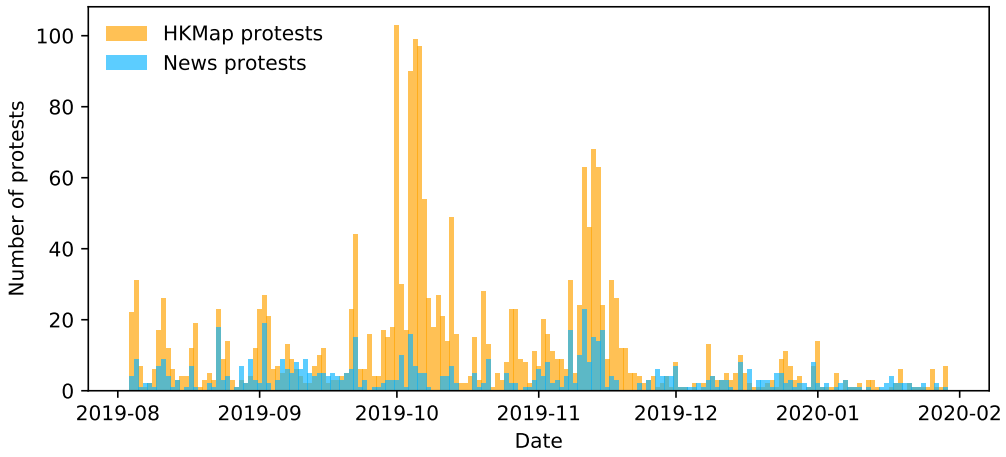

Figure 4. Daily number of unique protest locations identified in news sources (blue) and by clustering reports from HKMap.live (orange). The correlation between these time series is 0.5.

Another potential concern is our ability to measure arrests made by undercover agents. Given that undercover officers have the tendency to arrest protesters when they are alone and away from the main protest gathering sites, one may worry that these arrest events can hardly be observed. However, this issue is less concerning in the context of Hong Kong and the data we use, for several reasons. First, when undercover officers arrest lone or isolated protesters—such as individuals heading home or stepping away from the protest crowd to

[18] https://en.wikipedia.org/wiki/Timeline_of_the_2019-2020_Hong_Kong_protests
[19] https://antielabdata.jmsc.hku.hk/

eat or rest—it does not mean that no one can observe such arrests. Hong Kong is a small but densely populated city; many residents live near the streets where protests took place or pass by protest events regularly. Violent arrests often draw public attention and can be witnessed or recorded by nearby residents, passersby, or bystanders who are not involved in the protest. In fact, HKmap.live, a crowdsourced mapping service, was created to enable ordinary citizens to easily report incidents of police violence and their locations in real time using a mobile app on their smartphones. Given the widespread public support for the pro-democracy movement at the time, many Hong Kongese were willing to record and report such incidents. Second, during the protest waves, there were many Hong Kong citizens following and reporting events on the ground. They are often referred to as "citizen journalists" or "independent journalists" who report events on their independent digital platforms. They, often wearing yellow vests and carrying cameras, followed protesters and stood on the sidelines to document events unfolding.[20] They were frequently among the first to arrive at scenes of arrest and provided rapid, unfiltered coverage—often before mainstream media could respond. Appendix Figures A.2 to A.5 show examples of arrest footage reported by bystanders and passersby, and then shared on social media. The presence of citizen journalists, combined with voluntary reporting by ordinary citizens, allowed arrests carried out by undercover agents to be reported and recorded, even when they occurred in less crowded areas.

One may also worry that our unit of analysis might not be fine-grained enough to capture highly dynamic and localized arguments. Regarding spatial aggregation, we use district council constituencies as our spatial units. This choice is appropriate given Hong Kong's geography: it is a small, densely populated coastal city, with hilly and mountainous terrain. Only about one-quarter of its landmass is developed and densely inhabited, meaning most residents live in the flatter, central parts of the city. These developed and more habitable areas are further divided into 452 district council constituencies—compact units that typically contain only a few street blocks. As shown in Figures A.9 and A.10, these constituencies are often bounded by major roads and include only a few buildings and narrow lanes. They

[20]Figure A.6 shows footage of yellow-vested citizen journalists documenting arrests conducted by undercover police officers.

also represent the smallest administrative units commonly used in spatial analyses of Hong Kong (Yin, Huo and Lin, 2023; Chau and Wan, 2022). To ensure our findings are not driven by the choice of spatial scale, we also run our analysis using smaller grid cells (1km×1km). The results, presented in Appendix Tables A.7 and A.8, remain consistent.

Regarding temporal aggregation, we use four six-hour intervals—morning, afternoon, evening, and night—because these segments match human daily routines and reflect common patterns in protest behavior. Protest organizers typically avoid mobilizing during periods that are inconvenient for public gatherings, such as lunch (around 12 p.m.) and dinner (around 6 p.m.), as well as late at night when turnout is harder to generate. As a result, most protest events are scheduled within these four segments—particularly in the morning, afternoon, and early evening—and rarely extend beyond lunch or dinner hours. These four intervals therefore offer a reasonable and behaviorally grounded framework for analyzing protest activity. In terms of event duration, protests and marches often last about 3-4 hours in Hong Kong, excluding additional time needed for arrival and dispersal. Using a six-hour window allows us to capture the full trajectory of an assembly without running the risk of fragmenting a single protest event into multiple disjoint pieces. To ensure the robustness of our results, we also run our analysis using smaller three-hour time intervals. The results, shown in Appendix Table A.6, remain unchanged.

For our first hypothesis regarding distances, we conducted an additional check to validate our results. If it is true that the proximity of individuals to protest crowds influences the likelihood of their arrest by undercover police, then we should expect to see a null effect during periods when crowds are prohibited–essentially, when no treatment assignments are made–and distances to crowds should become irrelevant in determining arrests by undercover officers. Fortunately, our scraper continued to collect data in 2020, a period during which the Hong Kong government enforced stricter public gathering policies from mid-March and relaxed them starting in May. We utilized this Covid restriction period as a control group for a placebo test. The result in Appendix Table A.3 indicates that proximity to protest

crowds no longer has a clear effect on undercover police arrests, providing further support for our findings.

Turning to the second hypothesis about indoor spaces. One may also suspect that our result from the indoor spaces model may simply reflect that arrests by undercover police are more likely to happen in busy locations filled with transportation hubs and shopping malls. These arrests may have happened outside indoor spaces rather than inside. To address this concern, we use the Open Street Maps data to obtain shape files of all the indoor spaces where protesters tend to gather (e.g., the airport, the MTR stations, and shopping malls) in Hong Kong. We use the centers of these locations to create spatial bubbles (100m, 200m, 300m in radius) to approximate the indoor boundaries of these locations. Since these locations in Hong Kong typically do not go beyond 300 meters in radius, we can use it as the upper bound as a criterion to filter arrest data points located inside or outside these indoor spaces. Our result, shown in Appendix Table A.4, remains consistent across different radius specifications, lending additional confidence to our finding.

Lastly, if it is true that gatherings in confined indoor spaces tend to attract undercover police engaging in arrests, leading to arrests, then we should probably not observe the same pattern in big, open outdoor spaces. These locations often attract, and can accommodate, a large number of protesters. Larger crowds create disadvantages and increased risks for undercover officers if they take aggressive actions against the crowd. In Hong Kong, many reports have shown that large episodes of protest gatherings took place in outdoor parks and plazas during the 2019-2020 protest (Stokols, 2023; Kwok and Chan, 2022). Under this logic, we conduct a placebo test to examine the effect of outdoor parks on arrests by undercover officers. We sourced geolocation data for these parks from the Open Street Maps and used spatial bubbles to define park boundaries (100m, 500m, 1km in radius).[21] As shown in Appendix Table A.5, our analysis shows no consistent correlation between outdoor parks and arrests by undercover police. The placebo test provides additional support to our findings.

[21]All parks in Hong Kong are within a 1km radius.

Generalizability

Single-country studies often draw concerns regarding the generalizability of findings. However, considering the numerous reports of countries deploying undercover officers to police anti-government protests, we believe our findings possess broad relevance. The applicability of our findings is especially relevant for countries where governments, grappling with escalating protest violence, hesitate to execute overt mass crackdowns due to potential negative public perceptions and backlash. In such contexts, the discreet apprehension of radicalized individuals and leaders by undercover officers, especially in isolated areas to minimize confrontation, would become an appealing strategy for the law enforcement.

Systemic cross-national data on undercover policing are difficult to collect and do not yet exist, but individual cases from different countries offer additional support to our findings. For instance, during the 2020 Black Lives Matter protests in the US, reports highlight undercover federal officers using unmarked vehicles in Portland, Oregon, to arrest and detain protesters. The Acting Deputy Secretary even admitted to this tactic, justifying that "it was done to keep officers safe and away from crowds and to move detainees to a safe location for questioning."[22] Similarly, the 2019 Catalonia protests in Spain saw undercover officers, blending in as protesters, isolating and arresting individuals away from larger demonstrations. Such practices drew criticism, igniting concerns over police transparency and accountability.[23] Anecdotal evidence also suggests undercover policing in private shopping malls, such as the incident in Scottsdale, Arizona, during BLM protests where multiple arrests were made at Scottsdale Fashion Square mall after it was looted and vandalized. While it was not explicitly reported that undercover officers were involved, it was suspected that plainclothes police facilitated such operations.[24] As more data are collected and further research

[22] https://www.npr.org/2020/07/17/892277592/federal-officers-use-unmarked-vehicles-to-grab-protesters-in-portland

[23] https://www.elnacional.cat/en/politics/undercover-spanish-police-protest-supermarket_436050_102.html

[24] https://www.azcentral.com/story/news/local/scottsdale/2020/05/31/scottsdale-fashion-square-nearby-businesses-looted-and-vandalized/5299344002/

is undertaken, we believe that similar patterns concerning undercover policing, as we found in the context of Hong Kong, will emerge.

Conclusion

This study examines the strategy of deploying undercover policing during large gatherings. Drawing from unique micro-level data on protest policing in Hong Kong, we demonstrate that arrests made farther away from main protest crowds and in indoor spaces offer undercover agents tactical advantages in executing targeted arrests. It allows them to apprehend isolated protesters, minimizing the risk of violent clashes between the crowd of protesters and police. Taken together, our findings provide a rare look at the covert repression of protesters, pointing out a number of avenues for future research and offering policy suggestions.

One key implication of this study pertains to future research on state violence. Our work echoes growing calls by academics to pay more attention to covert (or less observed) repression. The mainstream literature on state violence often assumes that there is no coercion when states do not use violence. However, governments can maintain repression, often by concealing their suppressive tactics, making them more clandestine. Our research suggests that one important way to enrich this literature is to relax the current focus on overt suppression acts (such as torture, killings, and imprisonment) to engage in broader discussions about subtler, non-direct rights abuses. This includes covert strategies to curtail freedom of expression, freedom from fear, and the right to access public goods, among others. Although gathering data on these stealthy acts of repression may be more challenging, it offers the potential for devising more comprehensive policies to curtail human rights violations.

A second implication pertains to research on police violence and brutality. It is a pressing issue for many countries but most of the attention has been on violence committed by uniformed security forces, with little discussion on the force used by undercover or plainclothes agents (Ba et al., 2021; Shoub, Stauffer and Song, 2021).[25] Brutality is not exclusive to

[25]A different but related phenomenon is the government's use of non-state armed groups—such as pro-government vigilante groups or militia groups—to repress protests when the state seeks to avoid directly deploying official security forces and prefers to outsource repression in order to deny responsibility. While we do not observe systematic evidence of outsourced repression to pro-government militias in the context of

officers in uniform, and there is likely variation of violence committed by different law enforcement units that are worth exploring. In fact, undercover officers, who operate without badges and body cameras, may be more inclined to use excessive force due to their concealed identity. In New York, for instance, undercover and plainclothes officers constitute only a small portion of the NYPD but are implicated in nearly a third of the force's fatal shootings.[26] More work explicitly analyzing the practice of undercover law enforcement and their involvement in the use of force is thus necessary.

This study also concerns an important controversy related to covert repression that warrants further exploration. In our survey of undercover law enforcement within the context of Hong Kong and beyond, we find that it is not uncommon for undercover agents to be accused of provoking violent acts from protesters through so-called "decoy operations." Such tactics purportedly give state agents an excuse for making arrests. Within the repression literature, this strategy is identified as the use of "agent provocateurs" (Marx, 2013; Sullivan and Davenport, 2018). When undercover agents intentionally entice protesters into violent activities, it brings into question the legitimacy of arresting individuals who are misled into illicit actions, particularly if it is the law enforcement itself orchestrating these crimes. Our research offers initial insights into this pivotal issue but we need more research in this direction.

Hong Kong, future research could benefit from a comparative analysis of the use of violence by undercover police versus pro-government armed groups in other settings.

[26] https://theintercept.com/2018/05/09/saheed-vassell-nypd-plain-clothes/

References

Ba, Bocar A, Dean Knox, Jonathan Mummolo and Roman Rivera. 2021. "The role of officer race and gender in police-civilian interactions in Chicago." *Science* 371(6530):696–702.

Bagozzi, Benjamin E, Daniel Berliner and Ryan M Welch. 2021. "The diversity of repression: Measuring state repressive repertoires with events data." *Journal of Peace Research* 58(5):1126–1136.

Bass, Sandra. 2001. "Policing space, policing race: Social control imperatives and police discretionary decisions." *Social justice* 28(1 (83):156–176.

Bin Waheed, Umair, Saleh Al-Zahrani and Sherif M Hanafy. 2019. Machine learning algorithms for automatic velocity picking: K-means vs. DBSCAN. In *SEG International Exposition and Annual Meeting*. SEG p. D033S048R008.

Braga, Anthony A, Andrew V Papachristos and David M Hureau. 2014. "The effects of hot spots policing on crime: An updated systematic review and meta-analysis." *Justice quarterly* 31(4):633–663.

Butcher, Charles and Jonathan Pinckney. 2022. "Friday on my mind: Re-assessing the impact of protest size on government concessions." *Journal of Conflict Resolution* 66(7-8):1320–1355.

Carter, David B and Curtis S Signorino. 2010. "Back to the future: Modeling time dependence in binary data." *Political Analysis* 18(3):271–292.

Chau, Tak-Huen and Kin-Man Wan. 2022. "Pour (tear) gas on fire? Violent confrontations and anti-government backlash." *Political Science Research and Methods* pp. 1–11.

Choulis, Ioannis, Abel Escribà-Folch and Marius Mehrl. 2024. "Preventing Dissent: Secret Police and Protests in Dictatorships." *The Journal of Politics* 86(3):1104–1109.

Christensen, Darin. 2018. "The Geography of Repression in Africa." *Journal of Conflict Resolution* 62(7):1517–1543.

Conrad, Courtenay Ryals and Will H Moore. 2010. "What Stops the Torture?" *American Journal of Political Science* 54(2):459–476.

Cunningham, David. 2003. "Understanding state responses to left-versus right-wing threats: The FBI's repression of the new left and the Ku Klux Klan." *Social Science History* 27(3):327–370.

Davenport, Christian. 2007. "State Repression and Political Order." *Annual Review of Political Science* 10:1–23.

Davenport, Christian and Benjamin J Appel. 2022. "Stopping state repression: An examination of spells." *Journal of Peace Research* 59(5):633–647.

De Jaegher, Kris and Britta Hoyer. 2019. "Preemptive repression: Deterrence, backfiring, iron fists, and velvet gloves." *Journal of Conflict Resolution* 63(2):502–527.

Demirel-Pegg, Tijen and Karen Rasler. 2021. "The effects of selective and indiscriminate repression on the 2013 Gezi Park nonviolent resistance campaign." *Sociological Perspectives* 64(1):58–81.

Fijnaut, Cyrille and Gary T Marx. 1995. *Undercover police surveillance in comparative perspective*. Martinus Nijhoff Publishers.

Fong, Brian CH. 2022. "Diaspora formation and mobilisation: The emerging Hong Kong diaspora in the anti-extradition bill movement." *Nations and Nationalism* 28(3):1061–1079.

Hager, Anselm and Krzysztof Krakowski. 2022. "Does State Repression Spark Protests? Evidence from Secret Police Surveillance in Communist Poland." *American Political Science Review* pp. 1–16.

Hassan, Mai, Daniel Mattingly and Elizabeth R Nugent. 2022. "Political control." *Annual Review of Political Science* 25(1):155–174.

Hendrix, Cullen S and Idean Salehyan. 2017. "A house divided: Threat perception, military factionalism, and repression in Africa." *Journal of Conflict Resolution* 61(8):1653–1681.

Hill, Daniel W. and Zachary M. Jones. 2014. "An Empirical Evaluation of Explanations for State Repression." *American Political Science Review* 108(03):661–687.

Kaufman, Emily. 2016. "Policing mobilities through bio-spatial profiling in New York City." *Political Geography* 55:72–81.

Kwok, Chi and Ngai Keung Chan. 2022. "The making of contentious political space: The transformation of Hong Kong's Victoria Park." *Space and Culture* 25(4):615–632.

Liu, Howard. 2022. "Dissent Networks, State Repression, and Strategic Clemency for Defection." *Journal of Conflict Resolution* 66(7-8):1292–1319.

Liu, Howard, Ching-Hsuan Su and Yi-Ting Wang. 2025. "The Law or the Career? Autocratic Judiciaries, Strategic Sentencing, and Political Repression." *Comparative Political Studies* 58(10):2296–2328.

Liu, Howard and Christopher M. Sullivan. 2021. "And the Heat Goes On: Police Repression and the Modalities of Power." *Journal of Conflict Resolution* 65(10):1657–1679.

Lyall, J. 2009. "Does Indiscriminate Violence Incite Insurgent Attacks?" *Journal of Conflict Resolution* 53(3):331–362.

Marx, Gary T. 1988. *Undercover: police surveillance in America*. Univ of California Press.

Marx, Gary T. 2013. "Agents Provocateurs." *The Wiley-Blackwell Encyclopedia of Social and Political Movements* .

McCarthy, John and Clark McPhail. 2006. "Places of protest: The public forum in principle and practice." *Mobilization: An International Quarterly* 11(2):229–247.

McPhail, Clark and John D McCarthy. 2005. "Protest mobilization, protest repression, and their interaction." *Repression and mobilization* 21:3–32.

Nalepa, Monika and Grigore Pop-Eleches. 2022. "Authoritarian infiltration of organizations: causes and consequences." *The Journal of Politics* 84(2):861–873.

Olar, Roman-Gabriel. 2019. "Do they know something we don't? Diffusion of repression in authoritarian regimes." *Journal of Peace Research* 56(5):667–681.

Piotrowska, Barbara Maria. 2020. "The Price of Collaboration: How Authoritarian States Retain Control." *Comparative Political Studies* 53(13):2091–2117.

Ritter, Emily Hencken and Courtenay R Conrad. 2016. "Preventing and Responding to Dissent: The Observational Challenges of Explaining Strategic Repression." *American Political Science Review* 110(1):85–99.

Scharpf, Adam and Christian Gläßel. 2020. "Why Underachievers Dominate Secret Police Organizations: Evidence from Autocratic Argentina." *American Journal of Political Science* 64(4):791–806.

Schulte, Felix and Christoph V Steinert. 2023. "Repression and backlash protests: Why leader arrests backfire." *International Interactions* 49(1):1–30.

Shoub, Kelsey, Katelyn E Stauffer and Miyeon Song. 2021. "Do female officers police differently? Evidence from traffic stops." *American journal of political science* 65(3):755–769.

Slough, Tara and Christopher Fariss. 2021. "Misgovernance and human rights: The case of illegal detention without intent." *American Journal of Political Science* 65(1):148–165.

Sovey, Allison J and Donald P Green. 2011. "Instrumental variables estimation in political science: A readers' guide." *American Journal of Political Science* 55(1):188–200.

Starr, Amory, Luis Fernandez, Randall Amster, Lesley Wood and Manuel J Caro. 2008. "The impact of surveillance on the exercise of political rights: an interdisciplinary analysis 1998-2006." *Qualitative Sociology* 31(3):251–270.

Steinert, Christoph V and Christoph Dworschak. 2023. "Political imprisonment and protest mobilization: Evidence from the GDR." *Journal of Conflict Resolution* 67(7-8):1564–1591.

Steinert, Christoph Valentin. 2023. "The impact of domestic surveillance on political imprisonment: Evidence from the German Democratic Republic." *Journal of Conflict Resolution* 67(1):38–65.

Stokols, Andrew. 2023. "From the square to the shopping mall: new social media, state surveillance, and the evolving geographies of urban protest." *Urban Geography* 44(6):1166–1191.

Sullivan, Christopher M. 2016. "Political repression and the destruction of dissident organizations: Evidence from the archives of the Guatemalan national police." *World Politics* 68(4):645–676.

Sullivan, Christopher M and Christian Davenport. 2018. "Resistance is mobile: Dynamics of repression, challenger adaptation, and surveillance in US 'Red Squad'and black nationalist archives." *Journal of Peace Research* 55(2):175–189.

Teo, Elgar and King-wa Fu. 2021. "A Novel Systematic Approach of Constructing Protests Repertoires from Social Media: Comparing the Roles of Organizational and Non-organizational Actors in Social Movement." *Journal of Computational Social Science* 4(2):787–812.

Thomson, Henry. 2023. "The Bureaucratic Politics of Authoritarian Repression: Intra-Agency Reform and Surveillance Capacity in Communist Poland." *Political Science Research and Methods* pp. 1–16.

Uzonyi, Gary. 2022. "An Informational Theory of Genocide and Politicide During Civil War." *Comparative Political Studies* 55(6):933–962.

Weidmann, Nils B. 2016. "A closer look at reporting bias in conflict event data." *American Journal of Political Science* 60(1):206–218.

Weisburd, David and Anthony A Braga. 2006. "Hot spots policing as a model for police innovation." *Police innovation: Contrasting perspectives* pp. 225–244.

Yin, Weiwen, Weidong Huo and Danyang Lin. 2023. "The effects of state coercion on voting outcome in protest movements: a causal forest approach." *Political Science Research and Methods* 11(3):645–653.

Zhukov, Yuri M. 2015. "Population resettlement in war: Theory and evidence from Soviet archives." *Journal of Conflict Resolution* 59(7):1155–1185.

COVERT ASSIGNMENTS: UNDERCOVER INFILTRATION AND REPRESSION OF PROTESTS
ONLINE APPENDIX

Dependent Variables:	Protest zone	Close to protest zone	Protest zone	Close to protest zone
Model:	(1)	(2)	(3)	(4)
Weekend	0.003***	0.004***	0.003***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)
Day	0.000	0.002***	0.000	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Undercover Police	0.163***	0.013***	0.164***	0.013***
	(0.011)	(0.002)	(0.011)	(0.002)
Uniformed Police	0.032***	0.011***	0.032***	0.011***
	(0.002)	(0.000)	(0.002)	(0.000)
Arrests nearby	0.160***	0.201***	0.156***	0.196***
	(0.016)	(0.006)	(0.016)	(0.006)
Protest violence	0.038*	0.010***	0.038*	0.010***
	(0.022)	(0.001)	(0.022)	(0.001)
Fixed-effects				
CACODE	Yes	Yes	Yes	Yes
Fit statistics				
TimePoly	No	No	Yes	Yes
F-test (1st stage)	43.899	1,854.6	44.661	1,936.7

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

Table A.1. First Stage Results for 2SLS

	Logit	Logit
Close to protest zone	-1.898^{**}	-2.829^{***}
	(0.906)	(0.961)
Protest zone	1.765^{***}	1.775^{***}
	(0.131)	(0.132)
Undercover Police	25.872	26.203^{***}
	(2794.299)	(0.525)
Uniformed Police	0.674^{***}	0.697^{***}
	(0.254)	(0.256)
Arrests nearby	2.243^{***}	2.303^{***}
	(0.360)	(0.360)
Protest violence	0.516	0.473
	(0.349)	(0.353)
Indoor spaces	0.346^{**}	0.335^{**}
	(0.146)	(0.150)
LegCo	-0.254	-0.225
	(0.512)	(0.535)
Model	RE	RE
TimePoly	N	Y
Num. obs.	327248	327248
Num. groups: CACODE	452	452

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table A.2. Effects of Indoor Spaces (Logit)

	OLS	OLS	2SLS	2SLS
Closeness to protest zone	0.1343	0.1417		
	(0.1227)	(0.1250)		
Protest zone	0.1329***	0.1328***		
	(0.0480)	(0.0480)		
Undercover Police	0.0835***	0.0835***	0.0100	0.0095
	(0.0185)	(0.0184)	(0.0326)	(0.0317)
Uniformed Police	−0.0014**	−0.0014**	−0.0064**	−0.0065**
	(0.0007)	(0.0007)	(0.0031)	(0.0031)
Arrests nearby	0.2409***	0.2457***	0.0795	0.0805
	(0.0907)	(0.0922)	(0.2956)	(0.2252)
Predicted closeness to protest zone			−1.0871	−1.0962
			(2.1229)	(1.5801)
Predicted protest zone			1.5161**	1.5250**
			(0.7702)	(0.7285)
Model	FE	FE	FE	FE
TimePoly	N	Y	N	Y
Num. obs.	86784	86784	86784	86784
Num. groups: CACODE	452	452	452	452

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table A.3. Effects on Undercover Policing During Covid Lockdown

	300m	300m	200m	200m	100m	100m
Close to protest zone	-0.0421^{***}	-0.0426^{***}	-0.0253^{***}	-0.0256^{***}	-0.0089^{***}	-0.0089^{***}
	(0.0023)	(0.0023)	(0.0019)	(0.0019)	(0.0013)	(0.0013)
Protest zone	0.0315^{***}	0.0315^{***}	0.0232^{***}	0.0232^{***}	0.0116^{***}	0.0116^{***}
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0003)	(0.0003)
Undercover Police	0.0942^{***}	0.0942^{***}	0.0658^{***}	0.0658^{***}	0.0280^{***}	0.0280^{***}
	(0.0006)	(0.0006)	(0.0005)	(0.0005)	(0.0003)	(0.0003)
Uniformed Police	-0.0012^{***}	-0.0012^{***}	-0.0008^{***}	-0.0007^{***}	-0.0002^{**}	-0.0002^{**}
	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Arrests nearby	0.0318^{***}	0.0319^{***}	0.0163^{***}	0.0164^{***}	0.0014^{**}	0.0014^{**}
	(0.0012)	(0.0012)	(0.0010)	(0.0010)	(0.0007)	(0.0007)
Protest violence	0.0051^{***}	0.0051^{***}	0.0043^{***}	0.0043^{***}	0.0035^{***}	0.0035^{***}
	(0.0010)	(0.0010)	(0.0008)	(0.0008)	(0.0005)	(0.0005)
Indoor spaces	0.0005^{***}	0.0005^{***}	0.0005^{***}	0.0005^{***}	0.0003^{***}	0.0003^{***}
	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
LegCo	0.0032^{**}	0.0032^{**}	0.0025^{**}	0.0025^{**}	0.0010	0.0010
	(0.0014)	(0.0014)	(0.0012)	(0.0012)	(0.0008)	(0.0008)
Model	RE	RE	RE	RE	RE	RE
TimePoly	N	Y	N	Y	N	Y
Num. obs.	327248	327248	327248	327248	327248	327248
Num. groups: CACODE	452	452	452	452	452	452

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table A.4. Effects of Indoor Spaces on Arrests by Undercover Police (with radius)

	1000m	1000m	500m	500m	100m
Close to protest zone	-0.036^{***}	-0.037^{***}	-0.018^{***}	-0.018^{***}	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.000)
Protest zone	0.037^{***}	0.037^{***}	0.019^{***}	0.019^{***}	0.000^{**}
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Undercover Police	0.112^{***}	0.112^{***}	0.046^{***}	0.046^{***}	0.001^{***}
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Uniformed Police	-0.002^{***}	-0.002^{***}	-0.001^{***}	-0.001^{***}	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Arrests nearby	0.033^{***}	0.034^{***}	0.015^{***}	0.015^{***}	-0.000^{*}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Outdoor parks	0.000	0.000	0.000^{***}	0.000^{***}	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LegCo	0.014^{***}	0.014^{***}	0.023^{***}	0.023^{***}	0.001^{***}
	(0.002)	(0.002)	(0.001)	(0.001)	(0.000)
Model	RE	RE	RE	RE	RE
TimePoly	N	Y	N	Y	N
Num. obs.	327248	327248	327248	327248	327248
Num. groups: CACODE	452	452	452	452	452

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Note: 100m model with TimePoly can't fit due to limited variation in the outcome.

Table A.5. Effects of Outdoor Parks on Arrests by Undercover Police (with radius)

	OLS	OLS	2SLS	2SLS	OLS	OLS
Closeness to protest zone	-0.0462^{***}	-0.0467^{***}			-0.0467^{***}	-0.0471^{***}
	(0.0103)	(0.0103)			(0.0021)	(0.0021)
Protest zone	0.0459^{***}	0.0458^{***}			0.0461^{***}	0.0461^{***}
	(0.0065)	(0.0065)			(0.0005)	(0.0005)
Undercover Police	0.1572^{***}	0.1572^{***}	0.1222^{***}	0.1219^{***}	0.1575^{***}	0.1575^{***}
	(0.0103)	(0.0103)	(0.0145)	(0.0146)	(0.0005)	(0.0005)
Uniformed Police	-0.0014^{***}	-0.0014^{***}	-0.0079^{***}	-0.0079^{***}	-0.0013^{***}	-0.0013^{***}
	(0.0004)	(0.0004)	(0.0023)	(0.0023)	(0.0001)	(0.0001)
Arrests nearby	0.0251^{***}	0.0253^{***}	0.0206^{***}	0.0209^{***}	0.0251^{***}	0.0252^{***}
	(0.0067)	(0.0067)	(0.0068)	(0.0068)	(0.0008)	(0.0008)
Protest violence	0.0071	0.0071	0.0004	0.0004	0.0069^{***}	0.0069^{***}
	(0.0058)	(0.0058)	(0.0032)	(0.0032)	(0.0006)	(0.0006)
Predicted closeness to protest zone			-0.2039^{***}	-0.2057^{***}		
			(0.0559)	(0.0560)		
Predicted protest zone			0.3034^{***}	0.3054^{***}		
			(0.0847)	(0.0850)		
Indoor spaces					0.0003^{***}	0.0003^{***}
					(0.0001)	(0.0001)
LegCo					0.0022^{***}	0.0022^{***}
					(0.0008)	(0.0008)
Model	FE	FE	FE	FE	RE	RE
TimePoly	N	Y	N	Y	N	Y
Num. obs.	654496	654496	654496	654496	654496	654496
Num. groups: CACODE	452	452	452	452	452	452

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table A.6. Effects on Undercover Policing (3 hour window)

	OLS	OLS	2SLS	2SLS
Closeness to protest zone	-0.0865^{**}	-0.0862^{**}		
	(0.0344)	(0.0345)		
Protest zone	0.0670^{***}	0.0669^{***}		
	(0.0107)	(0.0107)		
Undercover Police	0.2062^{***}	0.2061^{***}	0.1362^{***}	0.1361^{***}
	(0.0156)	(0.0156)	(0.0224)	(0.0225)
Uniformed Police	-0.0035^{***}	-0.0035^{***}	-0.0162^{***}	-0.0162^{***}
	(0.0008)	(0.0008)	(0.0033)	(0.0034)
Arrests nearby	0.0883^{***}	0.0886^{***}	0.0375^{**}	0.0379^{**}
	(0.0212)	(0.0213)	(0.0169)	(0.0167)
Protest violence	0.0354^{***}	0.0354^{***}	-0.0219	-0.0220
	(0.0076)	(0.0076)	(0.0169)	(0.0169)
Predicted closeness to protest zone			-0.2847^{***}	-0.2864^{***}
			(0.0686)	(0.0684)
Predicted protest zone			0.5119^{***}	0.5130^{***}
			(0.1014)	(0.1017)
Model	FE	FE	FE	FE
TimePoly	N	Y	N	Y
Num. obs.	1225008	1225008	1225008	1225008
Num. groups: grid_id	1692	1692	1692	1692

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table A.7. Effects on Undercover Policing (1km grid, quarter-day level)

	Logit	Logit
Closeness to protest zone	2.269	0.068
	(3.206)	(4.848)
Protest zone	1.448^{***}	0.508^{**}
	(0.145)	(0.256)
Undercover Police	27.428	0.183
	(3402.432)	(0.274)
Uniformed Police	0.566^{**}	-1.220^{***}
	(0.286)	(0.139)
Arrests nearby	5.294^{***}	0.233
	(0.974)	(1.632)
Protest violence	0.797^{***}	0.347
	(0.200)	(0.289)
Indoor spaces	0.398^{*}	1.342^{***}
	(0.208)	(0.333)
LegCo	−0.336	0.752
	(0.753)	(5.823)
Model	RE	RE
TimePoly	N	Y
Num. obs.	1225008	1225008
Num. groups: grid_id	1692	1692

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table A.8. Effects on Undercover Policing (1km grid, quarter-day level, logit models)

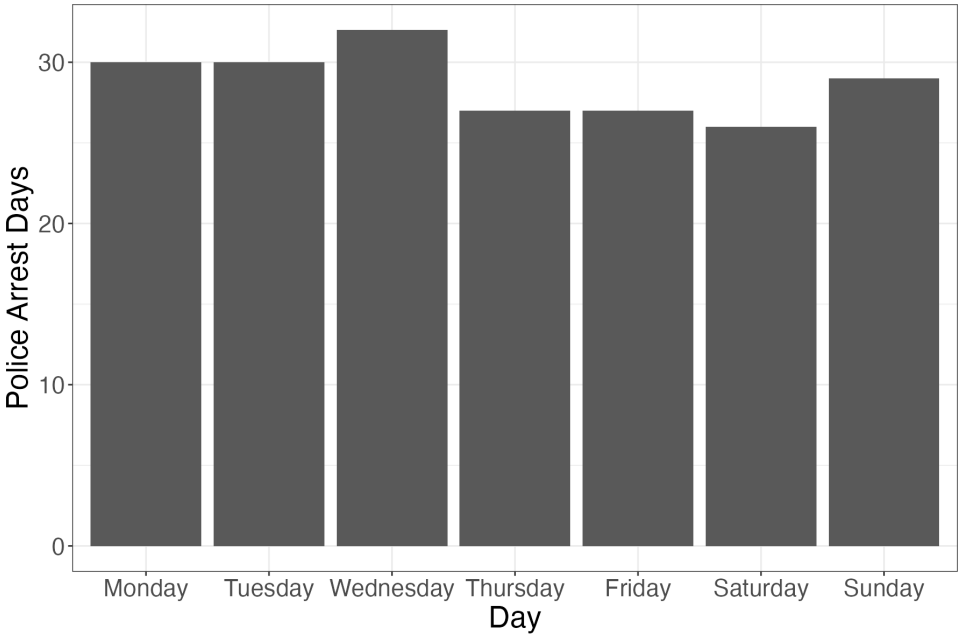

Figure A.1. Distribution of Police Arrests During the Day of Week (Data source: Hong Kong Security Bureau, from June 9 2019 to January 23, 2020)

Figure A.2. A few undercover police officers took down a lone protester they targeted, an incident captured by a citizen sitting inside a nearby restaurant. Source: X/Twitter. [Link]

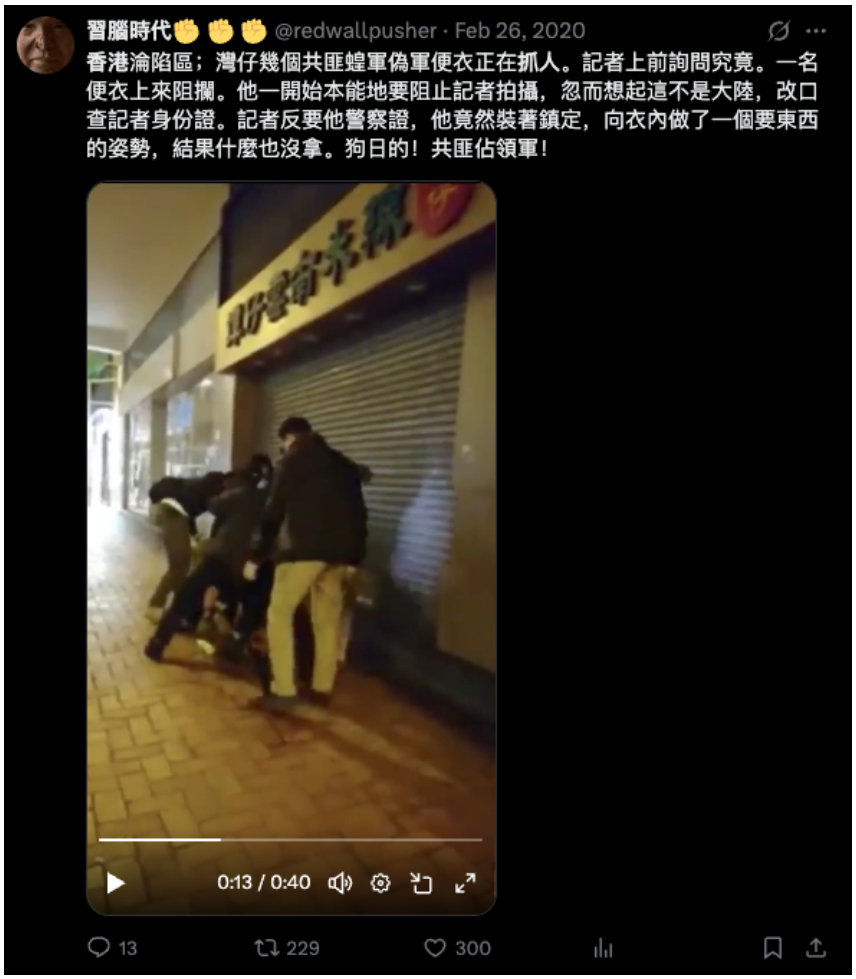

Figure A.3. A few undercover police officers took down a lone protester they targeted at a street corner, an incident captured by a citizen journalist. Source: X/Twitter. [Link]

Figure A.4. Undercover police officers took down a lone protester they targeted on the street, an incident captured and recorded by citizens nearby using their smartphones. Source: X/Twitter. [Link]

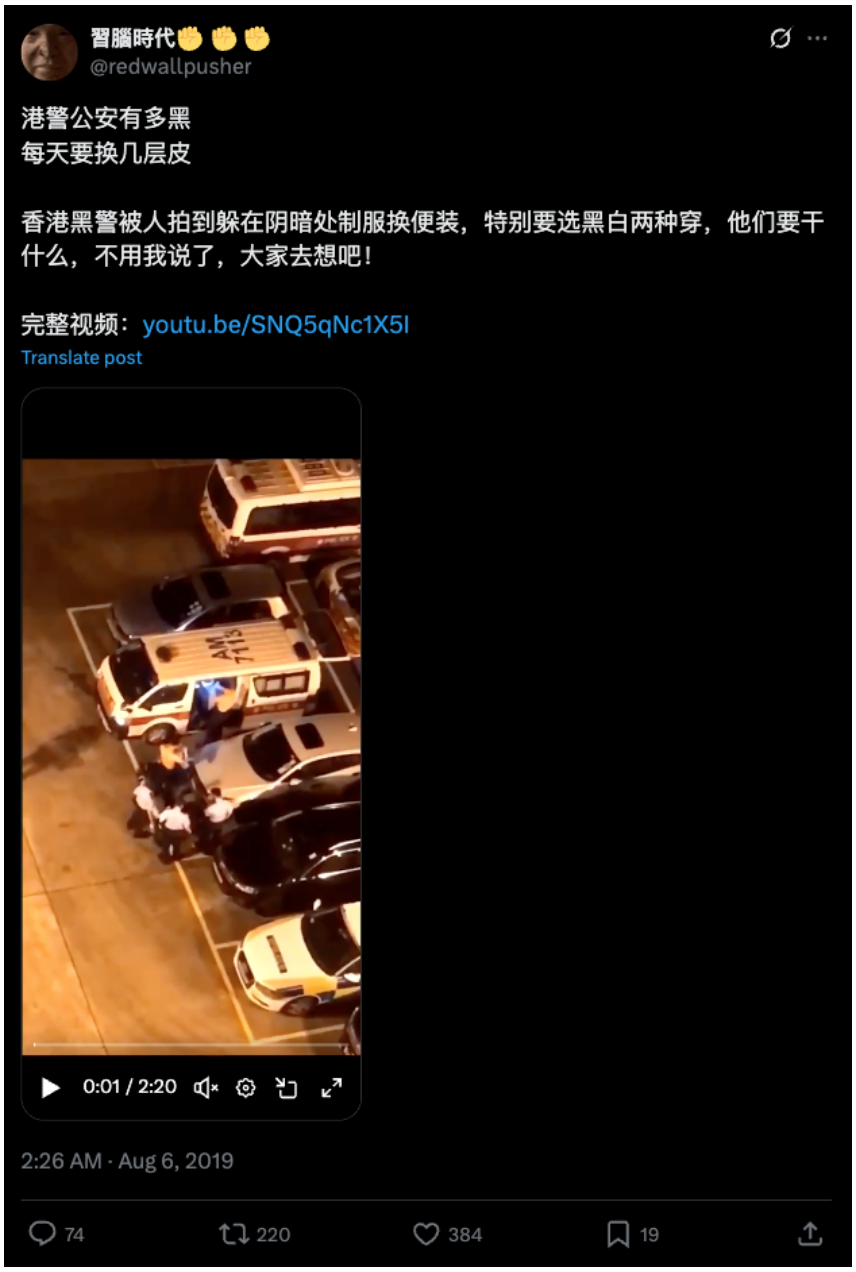

Figure A.5. A few police officers changed into civilian clothes to go undercover—an incident captured from above by a nearby resident recording from his/her home. Source: X/Twitter. [Link]

Figure A.6. An undercover arrest captured by citizen reporters wearing yellow vests (on the far right). Source: Hong Kong Free Press (HKFP)

Figure A.7. Call for A Protest Rally on January 12th (3-6pm) Source: Telegram

Figure A.8. Call for Marches and Rallies in August Source: Telegram

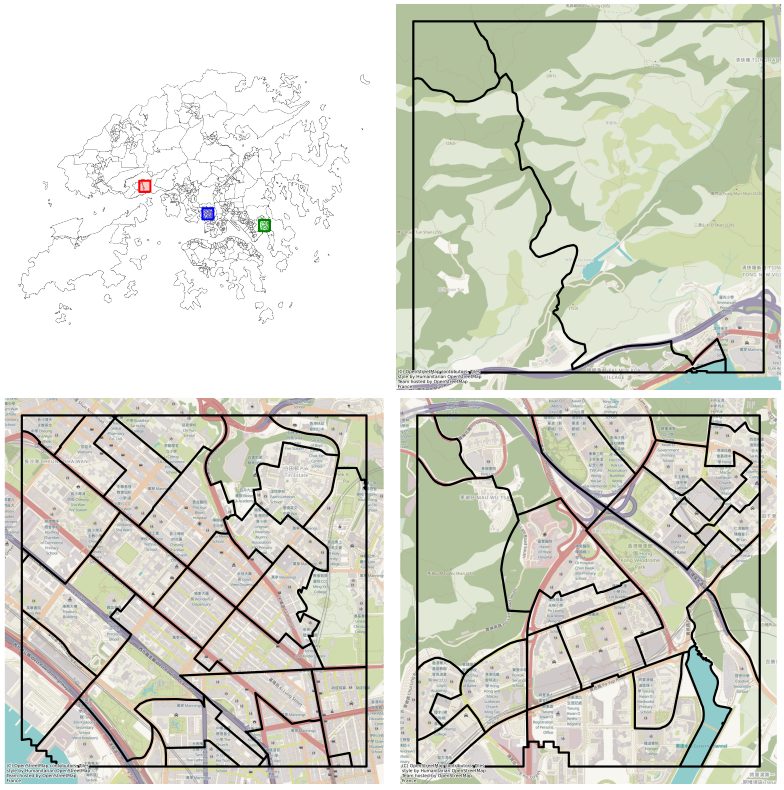

Figure A.9. Three close-up views of district council constituencies. Clockwise from upper left: A map of Hong Kong district council constituencies with three highlighted regions, a district constituency in a rural region (the red square), a district constituency in Kowloon (the green square), and a district constituency in Park Central (the blue square).

Figure A.10. A district constituency and its corresponding street view as shown on Google Maps Street View. The left panel displays one of the district constituencies (highlighted in the bottom-left panel of Figure A.9), and the right panel presents the street view of a narrow one-way lane at the location indicated by the blue arrow in the left panel.

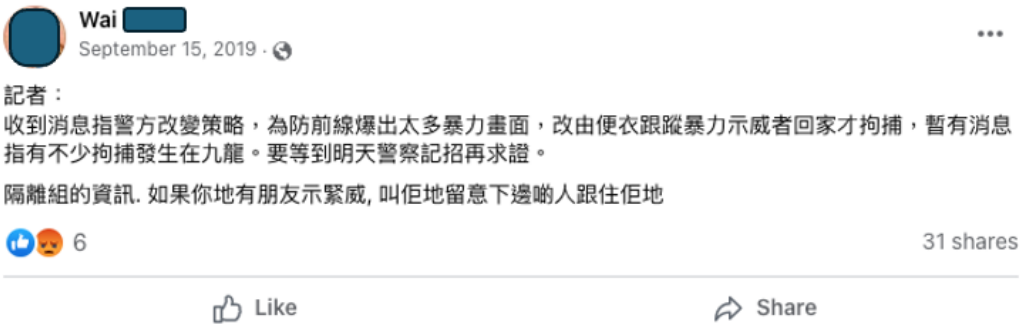

Figure A.11. Report of police changing tactics to follow radical protesters and arrest them as they made their way home *Note: Name and Facebook profile photo are partially obscured to protect identities.*

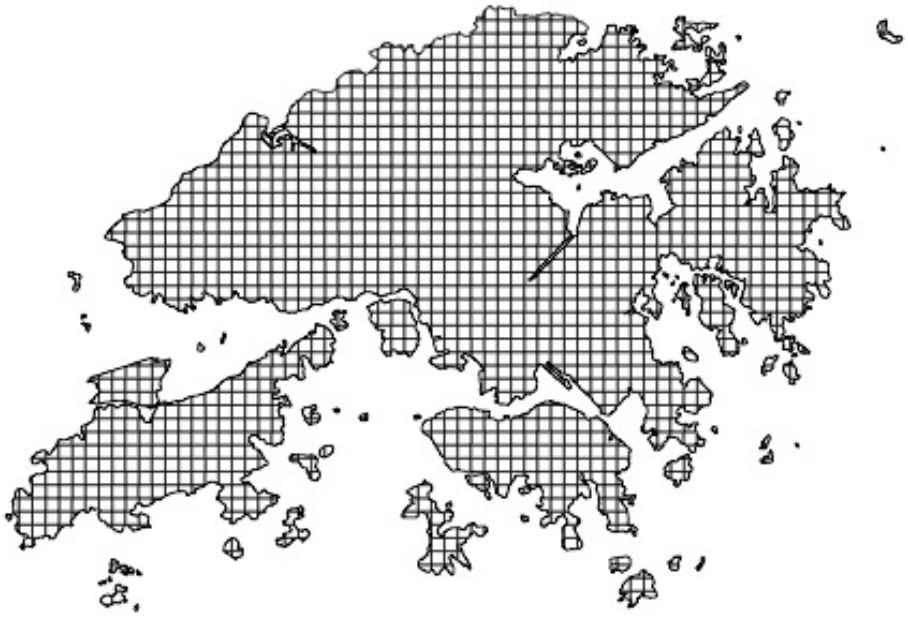

Figure A.12. The map of 1km×1km grid cells