Cover Stories

Michael F. Joseph
UCSD
mfjoseph@ucsd.edu

Matt Malis
Texas A&M University
malis@tamu.edu

September 3, 2023

Abstract

How do powerful states maintain plausible deniability for their secretive foreign interventions? Existing research focuses on the need for interveners to avoid direct exposure of their covert actions; this implies that they should refrain from public actions that attract attention and scrutiny while covert operations are underway. We uncover a countervailing logic whereby leaders enhance their plausible deniability by openly promoting public coercive actions and, even, inviting some scrutiny that raises the risk of exposure. Our insight is that foreign and domestic audiences draw strategic inferences. When they see a rival leader toppled in an unlikely coup, or an unsophisticated rebel group gain ground with advanced tactics, they infer—even in the absence of direct evidence—that covert action may have caused the unlikely outcome. Leaders can avoid these strategic inferences through a *cover story*: a plausible alternative story that explains how they got the outcome they wanted without resorting to controversial covert action. We apply our theory to Operation PB-SUCCESS, the CIA operation that covertly overthrew democratically elected Guatemalan president Jacobo Arbenz. We show that Eisenhower was concerned about strategic inferences. He pursued public actions—in the form of an arms embargo, and a diplomatic campaign through the Organization of American States—as a cover story to maintain plausible deniability both during the covert operation and after its successful completion. The theory advances our understanding of how leaders evade accountability for behaviors that violate norms and laws across a wide range of foreign and domestic policy domains.

In March 1960, the CIA began organizing Cuban exiles to oust Fidel Castro. Eisenhower demanded that the CIA took extraordinary precautions to avoid direct evidence of U.S. involvement (Poznansky, 2020). But as CIA agents were quietly meeting Cuban contacts and building training remote, secretive bases in Guatemala and Florida, Eisenhower initiated a public show-down with Castro. In December 1960, Eisenhower publicly announced a complete elimination of Cuba's sugar import quota, justified by Cuba's "deliberate hostility" towards the U.S. and increasing economic integration with the the Soviet bloc.[1] The next month, the administration formally severed diplomatic ties with Cuba[2]—a wholly symbolic gesture, as the U.S. ambassador had already been recalled and communication between the governments already ceased entirely.[3] Shortly after, New York Times reports began speculating that the CIA could be training and equipping an invasion force.[4][5]

Why would Eisenhower invite scrutiny on an issue he wanted to keep secret? The conventional wisdom dictates that he wouldn't. Scrutiny raises the risk that targets or the media uncover direct evidence of secret policies (Krcmaric, 2019). Therefore, states should avoid attracting attention as covert operations are ongoing (Joseph and Poznansky, 2018). The desire to avoid scrutiny is important, but it is not all that decision-makers consider. We uncover a countervailing incentive, whereby states, counterintuitively, pursue overt actions to plausibly deny their covert actions.

We arrive at our insight through a different conceptualization of plausible deniability (Poznansky, 2022). Consistent with existing research into World Order and covert action, we focus on a setting where authorizing certain military actions, perhaps because actions violate important norms or international laws, will generate audience costs (Ikenberry, 2000; Goddard, 2018; Bull, 2002). Thus, powerful states (we label them Interveners) exploit covert action to achieve their objectives, and deny that they did so. Existing research explores how Interveners avoid direct evidence of their involvement (eg Carnegie and Carson, 2018; Poznansky, 2019).[6] We argue that international and domestic audiences do not only rely on direct evidence to determine whether or not a covert action has taken place. Rather, they draw strategic inferences about whether a state authorized a covert

[1] https://www.presidency.ucsb.edu/documents/statement-the-president-upon-issuing-proclamation-fixing-the-cuban-sugar-quota-z
[2] https://history.state.gov/countries/cuba
[3] https://adst.org/Readers/Cuba.pdf, p.53-59
[4] NY Times, Jan 6 and Jan 14, 1961
[5] We'll never know how much Eisenhower's public actions caused reporters to investigate broader US-Cuba activities. But they likely did not help.
[6] See Carnegie (2021) for a review.

action, given the outcomes they observe, *and* their understanding of the Intervener's incentives and capabilities. When international audiences observe an outcome that they know the Intervener wanted, and that was unlikely to have come about absent foreign intervention, they can infer—even without direct evidence—that the Intervener likely exploited covert action to achieve this outcome.

If audiences draw strategic inferences, then plausible deniability is vastly more complicated than past secrecy scholars have recognized (see Carnegie, 2021). This affects all coercive practices where attribution is at issue (Baliga, Bueno de Mesquita, and Wolitzky, 2020), including covert actions (Spaniel and Poznansky, 2018), grey-zone conflict (Schram, 2022), rogue state management (Coe, 2018), cyber-conflict (Axelrod and Iliev, 2014) and election meddling (Levin, 2021). For instance, how can the US avoid strategic attribution for the STUXNET cyber-attack when Iran's nuclear research facilities do not just fail on their own (Lindsay, 2013)? How can Iran disclaim responsibility when uneducated terrorists develop technologically sophisticated weaponry—such as explosively formed penetrators (EFPs)—to penetrate US armored vehicles in Iraq?[7]

We argue that interveners can utilize overt policies to generate a *cover story*: that is, a plausible alternative story that can explain how the Intervener achieved the outcome she wanted, without having resorted to covert action that violate international laws and norms. Before a covert action has succeeded, public statements and actions that draw attention to the issue raise the risk of direct exposure. But after the covert action succeeds, those same public actions serve to shape observers' retrospective evaluations in a way that places the Intervener in a more favorable light: upon observing a successful policy outcome, the audience is more willing to attribute that outcome to the accompanying public action, and less likely to suspect that the Intervener engaged in morally repugnant covert actions to achieve the outcome that the audience observed.

To be clear, we find that both taking public actions to generate a cover story, and avoiding public actions to avoid scrutiny, can (under different circumstances) constitute rational mechanisms to sustain plausible deniability. Our analysis identifies clear conditions under which one is better than the other. Cover stories are best when outside observers believe that a foreign policy outcome is unlikely to come about by random chance (i.e. absent foreign intervention); when the level of transparency surrounding covert interventions is low enough as to render the absence of direct evidence uninformative; and when public actions invite some, but not too much, additional scrutiny.

[7] https://www.govinfo.gov/content/pkg/CHRG-112shrg76693/html/CHRG-112shrg76693.htm

Further analysis reveals a more nuanced picture of how the cover story mechanism operates. When Interveners use cover stories too frequently, they become less effective as a tool for reputational management. This highlights that Intervieners must employ the tactic sparingly so as to maintain unpredictability. Most surprisingly, we find that the frequency that Intervieners should deploy cover stories can be non-monotonic in the degree of heightened media scrutiny that their public actions invite. Increased scrutiny can make cover stories more attractive for the Intervener because if scrutiny does not unearth covert action, the audience is more confident that no covert action took place as their scrutiny is more effective. This logic applies up to a point where the risk of direct exposure from scrutiny becomes so great that it discourages cover stories entirely.

We illustrate our mechanism with evidence from Operation PBSUCCESS, Eisenhower's covert intervention to oust the Guatemalan president in 1954. It is well known that administration officials feared international backlash, and therefore only considered the mission successful if plausible deniability was achieved (Schmitz, 1999). Standard accounts show that the Dulles brothers and CIA sought to avoid direct evidence of US involvement through tight operational controls, or distancing themselves publicly from the coup plotters as the coup was ongoing. Our theory illuminates other dynamics. We find direct evidence that the Administration and CIA planners believed that many in Latin America would blame the US even absent direct evidence because of strategic inferences. We provide evidence that highly publicized shipping embargoes, and sanctions and protests registered through the OAS were partly designed to disclaim responsibility for the coup plot. We further show that after the mission was complete, the US government refers to their public actions at the OAS as a cover story to disclaim covert action.

Our theory holds direct implications for theories of conflict where attribution is ambiguous. Outside of international security, it holds implications for research into international norms, laws and world order (Lake, Martin, and Risse, 2021; Farrell and Newman, 2021). There is mounting evidence that violating international laws and norms is costly (Huth, Croco, and Appel, 2011; Terman and Byun, 2022). But critics still worry that constraining effects are limited because powerful states can exploit covert action (Carson, 2018). Our theory suggests a practical limit on how frequently states can use covert action to circumvent international responsibilities. If every leader in East Asia that refused a Chinese military base died of a heart attack, we would infer that China was behind at least some of them. Outside of international relations, our model will interest

scholars of electoral accountability and adverse selection (Ashworth, 2012). This research has rationalized seemingly bizzare behaviors including showing-off (Gleason, 2017), admitting ignorance (Backus and Little, 2020), and extreme ideological policies (Izzo, 2022). In a similar vein, we rationalize why leaders implement, then broadly publicize, ineffective and costly policies. Finally, we offer a rational basis for persistently held conspiratorial beliefs studied in comparative and American politics (Nyhan and Zeitzoff, 2018; Oliver and Wood, 2014; Uscinski, Klofstad, and Atkinson, 2016).

1 Secrecy, international reputation and plausible deniability

In this section, we advance a conceptual innovation for research into secrecy. In the next section, this will justify our use of a principal-agent model, and the technical innovations we make to the standard formulation. First, we review a common arguments that states who consider coercive actions fear backlash when their coercive actions are unscrupulous. We define unscrupulous actions (and leaders) in more detail below. Second, we argue that existing research narrowly construes plausible deniability. We argue that much insight can be gained into research on World Order, secrecy, conspiratorial beliefs, and patterns of intervention more broadly from a strategic definition of plausible deniability.

1.1 Why take covert action? The international reputation approach

Perhaps the main reason that leaders authorize secret coercive actions is to avoid the reactions from various audiences who would find a coercive action unpalatable. Interveners are worried that the Target will retaliate (Carson, 2015). Interveners are also often concerned about the reactions of third-party audiences[8] when coercive actions violate international norms and laws (Goddard, 2018; Owen, 1994; Colgan, 2021; Terman and Byun, 2022). As a result, states often deploy coercive actions that violate international laws and norms in secret to avoid reputational costs (Poznansky, 2019, 2020; Bull, 2002; Morse and Pratt, 2022).

These empirical studies supply evidence that third-party audiences harm powerful states for undesirable coercive actions. But in each study, who the audience is, and the conditions under

[8]When we say audience, we mean actors other than the government of the Target of a coercive action.

which they are willing to impose harm is context-dependent. It depends, in part, on the norms and laws to which the Interviener subscribes (Lake et al., 2021). For example, democracies that want to coerce fellow democracies, or meddle in another democracy domestic affairs, turn to secrecy to avoid a bad international reputation (Downes and Lilley, 2010; Reiter and Stam, 2002). States that want to commit human rights atrocities, or support an ally commit atrocities often do it in secret because they are worried about backlash from the international community (Krcmaric, 2019). It also depends on the norms that a specific audience cares about. For example, domestic audiences punish leaders for coercive foreign policies that contradict the Interveiner's specific national values (Downs, 1964). Whereas regional actors, such as the members of the African Union, may care if members violate sovereignty norms, no matter what their specific national values are.

It also depends, in part, on the extent to which laws and norms conflict (Farrell and Newman, 2021), and therefore the extent that a state can craft a justification (Stein, 2000). For example, the decision to intervene against state-sponsored genocide pits a US commitment to uphold sovereignty against human rights. In some cases, international audiences want powerful states to apply (avoid) coercive power even though it violates certain laws and norms (Finnemore, 2003). In other cases, the international community may express outrage even if coercive actions are technically justifiable under international law.

The important point for our strategic theory is that in some contexts, some coercive actions invite severe backlash but others will not. For brevity,[9] we refer to coercive actions that an Intervene worries will cause backlash as **unscrupulous** actions. For example, concerns about the United States' reputation in Latin America, and the response of regional partners, in part, motivated Congress to prohibit Reagan from arming the Contras to oust Nicaraguan president Daniel Ortega. But Congress was less opposed to sanctions, or diplomatic isolation to influence Nicaraguan policy. Since the backlash is confined to arming the Contras, and other efforts to overthrow Ortega through direct military action, it makes sense that members of Reagan's administration believed they would face fierce backlash if they publicly took these actions. This drove them to covert action. Our theory starts at the point where Reagan must chose a policy outcome, taking these reactions to different policy tools as given.

[9]Given the normative debates just described, we deliberately avoid the terms ethical, compliant with international laws.

To be clear, our theory accounts for policy scenarios that vary in how unscrupulous a policy response is because we parameterize the direct costs of covert and public actions, as well as the costs that audiences impose on Interveners. The scrupulous action in the Contra case would be diplomatic isolation with smaller costs (0 is the minimum bound on our cost parameters). The unscrupulous action could be intervention, with larger costs (there is no maximum backlash cost in the model). We discuss variation in the costs and the strategic logic it generates.

1.2 Plausible deniability: A strategic theory

Like existing research into the strategic logic of covert action, we believe that if a leader predicts that if an action is scrupulous, then a leader need not pursue it in secret. We also agree that if the audience is likely to learn of the leader's unscrupulous covert action then the Intervener avoids taking that action (Joseph and Poznansky, 2018).

Where we depart from existing research is in what it takes to sustain plausible deniability. In the mechanics of existing theories, whether states sustain plausible deniability depends on whether a random variable exposes their action (Spaniel and Poznansky, 2018). Substantively, this represents a farmer that stumbles upon a CIA training facility, or an enterprising reporter that photographs operatives during a mission. This conceptualization focuses on the suppression of *direct* evidence of covert operations. Although they do not describe it that way, the focus on direct evidence is far-reaching. In a comprehensive review, Poznansky (2022, 523-524) identifies only three "threats to plausible deniability" at the state level: leaks, rival intelligence, or electronic recording—all variants of direct evidence.

We argue that audiences are clever. This creates a strategic barrier for sustaining plausible deniability that is not explored in existing studies. Specifically, audiences draw inferences from the strategic context. This includes their knowledge of the powerful state's preferred policy outcome, and their expectations about whether those outcomes would occur if the powerful state did not intervene. For example, in the late 1980s, Iranian dissidents living in Europe were 100,000 times more likely to be murdered in a robbery gone bad than the average European citizen. There was no direct evidence that Iran sponsored these murders. And yet, the German government tried the IRCG in absentia for their actions (Hakakian, 2011). Indeed, it is fear of strategic inferences that drive powerful states to create incredibly elaborate covert actions. For example, the Soviets

developed undetectable poisons that presented as heart attacks so that no one would know that they assassinated dissidents. The United States experimented with building Tsunami-generators to destroy cities precisely because they knew that no one would suspect that this technology was viable, and therefore they could disclaim responsibility for it (Houghton, 2019).

If the goal of plausible deniability is to avoid backlash for unscrupulous policies, then mission success requires that Interveners avoid these costs by convincing relevant audiences, to a sufficiently high level of confidence, that they were not responsible for those policies. As the mission planning is happening, Interveners must avoid direct evidence of their involvement. After the mission is complete, they must avoid strategic inferences of their culpability.

The challenge of strategic inferences intuitively leads to instant explanatory gains for important cases. Strategic inferences help clarify why states would use different covers under different conditions. For example, it partly clarifies why the CIA they continue to apply the Mosaddeq model for regime change while their sponsorship is unknown, but switch methods once the Bay of Pigs was exposed.

Strategic inferences also rationalize certain incorrect attributions and even conspiratorial beliefs. In 1950, the communist-leaning Bulgarian government foiled a coup plot. Given the US position on communism, and the Ambassador's broad personal relations across Bulgaria, the Bulgarian government inferred that US Ambassador, Donald Heath, was involved. They declared Heath persona non-grata, leading Bulgarian-US relations to sever.[10] With 70 years of hindsight, there is no evidence that the US was involved.

In the modern world, conspiracy theories often grip mass support inside America or across a particular region of the world. For example, it is widely held across the Middle East that the US orchestrated 9/11 attacks in secret to justify a military intervention in the Middle East (Nyhan and Zeitzoff, 2018). Partly, in response to this belief, "Osama bin Laden was among the top three leaders most often trusted to "do the right thing" by survey respondents" (Gentzkow and Shapiro, 2004). Scholars explain beliefs like these through psychological processes (Oliver and Wood, 2014), and a combination of identity and elite cues (Uscinski et al., 2016). But our theory suggests a potential rationalist motivation for them. These beliefs can be rationalized as follows: the US is heavily fortified, rendering a 9/11 attack unlikely; the US government wants to intervene, but

[10] https://bg.usembassy.gov/our-relationship/policy-history/io/

needs a pre-text for it; and the US has a history of complex covert actions. In fact, qualitative surveys show that Middle Eastern subjects frequently point to the 1953 CIA-led coup plot against Mossadeq to support the plausibility of their 9/11 conspiritorial beliefs (Kinzer, 2003). Even in the absence of direct evidence, strategic inferences can lead rational observers to hold beliefs that would otherwise appear conspiratorial.

There is only so much we can claim from an intuitive account. It is not instantly clear how leaders navigate the trade-offs between avoiding direct evidence during mission planning, and building a long-term cover story. The strategic problem is especially complicated because leaders know that audiences are drawing strategic inferences and will adjust their behavior accordingly. Audiences also know that leaders are strategic and are trying to manipulate their beliefs. If leaders know others will infer they are responsible for outcomes, they may not take covert action. But if leaders know others will assume they are responsible, they may as well take the action anyway. The audiences also know that leaders are weighing their responses and must think through the leader's strategic incentives to avoid covert action. We are not smart enough to reason through this complicated problem intuitively and we see good reasons that strategic problems could encourage and discourage covert actions, and therefore raise or reduce suspicion that follows from strategic inferences.

2 A principal-agent theory of unscrupulous covert actions

We utilize a formal principal-agent model to analyze the decision to pursue an unscrupulous covert action. When applied to the study of politics, principal-agent models are often used to examine the relationship between an "agent" (e.g. a political leader or other elected official) who sets policy on behalf of a "principal" (e.g. a domestic voter), who holds the agent accountable for his performance (see Miller, 2005). Within this context, the strategic challenge that interests us is adverse selection (Ashworth, 2012). That is, the agent's value to the principal depends on their level of quality along a privately-known attribute. The principal's task is to discern the agent's quality, retain the high-quality agents and replace the low-quality agents. Researchers have examined many different attributes including competence, honesty, political ideology. When applying these models to the study of foreign policy, scholars often consider the leader's privately-known type to be either

his competence in executing international conflicts (Smith, 1998; Ramsay, 2004), or his ideological alignment with a representative voter (Schultz, 2005; Fang, 2008; Malis, 2023).

We know from past principal-agent theories that strategic inferences by principals can vastly complicate the incentives that agents face. These dynamics can explain a wide range of policymaking behavior, such as "pandering"—taking popular actions that the leader privately believes to be contrary to the public interest—and, more puzzlingly, "fake leadership"—taking *unpopular* actions that are against the public interest, in order to convincingly demonstrate that one is not pandering (Canes-Wrone, Herron, and Shotts, 2001; Maskin and Tirole, 2004). These incentives can distort policymaking to the point that voters may actually be better off receiving *less* information with which they can hold their leaders accountable (Ashworth and Bueno de Mesquita, 2014). Previous work has also shown that the nature of the reputational concern (e.g., whether it pertains to preferences or competence) has important implications for the agent's behavior, and ultimately the principal's welfare (Fox and Shotts, 2009).

Thus, we purpose-built our principal-agent model to meet the novel reputational concern described above. That is, we focus on the leader of the intervening state's *scruples*. Loosely, this is the extent to which a leader has internalized the international norms and institutional commits that her state commits to uphold, and her willingness to violate them when it is politically expedient to do so.[11] Holding fixed the nature of the foreign policy issue, the feasibility of different policy issues, and the risk of audience backlash, we say that scrupulous and unscrupulous leaders differ in the intrinsic value they place on adhering to international laws and norms, or promises made. We assume that the leader seeks approval from some audience (foreign, domestic, or both) who seeks to "reward" scrupulous leaders (at the ballot box, or through future cooperation at the international level) and "punish" leaders they believe to be unscrupulous.

2.1 Technical Set-up

Our model analyzes a strategic interaction between a leader L of an Intervener state, and an audience A who holds the leader accountable for his policy actions and outcomes. L can represent

[11] As reviewed extensively in the concepts section, there are more nuanced normative debates that illuminate exceptions to both the normative and legal components. Generally, we mean that the leader is a type that has internalized the values that would drive the leader to behave the way that the international community wants them to behave.

Figure 1: Game Sequence

1. L's type $\theta \in \{0,1\}$ is realized and observed privately. A holds prior belief $Pr(\theta = 1) = \pi \in (0,1)$.
2. Policy feasibility $\omega = (\omega_c, \omega_p) \in \{0,1\}^2$ is realized and observed privately by L. A holds prior belief $Pr(\omega_j = 1) = \tau_j \in (0,1)$ for $j = c, p$.
3. L chooses public action $a_p \in \{0,1\}$, which A observes, and covert action $a_c \in \{0,1\}$, which A does not observe directly.
4. Policy outcome $y \in \{0,1\}$ is realized, according to the probabilities given in (1).
5. Covert revelation $z \in \{0,1\}$ is realized, with $Pr(z = 1|a) = a_c(\lambda + a_p\delta)$.
6. A observes $(a_p, y, z) \in \{0,1\}^3$ and chooses $r \in \{0,1\}$.

the leader acting alone, or in concert with his foreign policy advisory team, insofar as the latter's political and policy objectives align with those of the leader. As discussed above, the audience can represent voters, mass publics or political elites across different foreign countries—including the target of L's intervention, adversary states, or third-party states.

The sequence of moves and information structure of the game is reported in Figure 1. The leader has two policy levers available. He can enact either one, both, or neither. First is a public action $a_p \in \{0,1\}$, which is taken openly and is understood to be in compliance with commonly-supported international norms and institutions. Second is a covert action $a_c \in \{0,1\}$, which is taken secretly and is understood to violate those norms. Referring back to our opening anecdote for concreteness, a_p can represent the Eisenhower administration's imposition of economic pressure on Cuba through the slashing of sugar quotas, while a_c can represent the various attempts made to oust or assassinate Castro through CIA-supported Cuban exiles or through agents operating secretly within the country.

Each policy is either feasible or infeasible, which we denote as $\omega_j \in \{0,1\}$ for $j = c, p$. When a feasible policy action is taken, it leads probabilistically to a policy success or failure, $y \in \{0,1\}$;

that is,

$$Pr(y=1|\omega,a)=\begin{cases}\alpha_p, & a_p\omega_p=1\\ \alpha_c, & a_c\omega_c=1\,\&\,a_p\omega_p=0\\ \alpha_0, & a_c\omega_c=0\,\&\,a_p\omega_p=0\end{cases} \tag{1}$$

where α_0 denotes the baseline probability of success absent any intervention from L, or due to "random luck". We assume $\alpha_0 < \alpha_c < \alpha_p < 1$, meaning that both actions increase the odds of policy success but do not guarantee it; the assumption that $\alpha_p > \alpha_c$ means L has better information about the effectiveness of public action as compared to covert action.[12]

There are two aspects of the game which are privately known by the leader, and unobserved by the audience. The first pertains to the feasibility of each policy lever. L observes both ω_c and ω_p, while A holds prior beliefs $Pr(\omega_j = 1) = \tau_j$. Thus the ratio $\frac{\alpha_j}{\tau_j}$ can be loosely thought of as representing the degree of informational asymmetry between L and A regarding the effectiveness action j: in the limiting case of $\tau_j \to 1$, the leader and audience are equally well-informed; whereas a high α_j and low τ_j indicate that the leader is much better informed than the audience is regarding the likelihood that action j will lead to a policy success.

The leader's second informational advantage over the audience pertains to the leader's intrinsic willingness to abide by international norms and institutions. We refer to this quality as the leader's *scruples*, and denote it formally as a binary variable $\theta \in \{0, 1\}$, with $\theta = 1$ denoting a scrupulous leader. L knows his own type θ, while A holds a prior belief, $Pr(\theta = 1) = \pi \in (0, 1)$, which she can revise over the course of the game. Unscrupulous leaders may be induced, through strategic and reputational concerns, to refrain from illegal, or norm-violating foreign policy behavior; but scrupulous leaders are intrinsically unwilling to employ such tactics.

After the leader takes his action and the policy outcomes are realized, the audience chooses whether to reward or punish the leader, $r \in \{0, 1\}$. Interpreting A as a political leader of a friendly third country, for instance, we can think of "rewarding" L ($r = 1$) as maintaining cooperation on future foreign policy issues, while "punishing" ($r = 0$) would involve defecting from L's bloc or otherwise defying L's leadership on the global stage. Rewarding L brings A a payoff normalized to 1 if L is scrupulous, and 0 if L is unscrupulous, while punishing brings A a payoff of $\bar{\mu} \in (0, \pi)$:

[12] Note that this assumption does not rule out the possibility covert action being more effective overall than public action, if τ_c is sufficiently higher than τ_p.

that is,

$$U_A(r) = r\theta + (1-r)\bar{\mu} \tag{2}$$

The audience observes three pieces of information, on which they can base their decision to reward or punish. First, A observes whether L took the public action a_p. Second, A observes the policy outcome $y \in \{0,1\}$. Third, A probabilistically observes the revelation of L's covert action a_c. Specifically, let $z \in \{0,1\}$ denote whether covert action is revealed, with

$$Pr(z=1|a) = \begin{cases} 0, & a_c = 0 \\ \lambda, & a_c = 1 \,\&\, a_p = 0 \\ \lambda + \delta, & a_c = 1 \,\&\, a_p = 1 \end{cases}$$

Whenever the leader refrains from covert action, A observes $z = 0$; but if the leader does take covert action, A observes $z = 1$ with probability $\lambda \in (0,1)$, or with probability $\lambda + \delta$ if the leader takes public action in addition to covert action. Thus λ denotes the baseline risk of revelation, while δ denotes the extent to which that risk is elevated by L's drawing attention to the issue through public action. Altogether, A observes $(a_p, y, z) \in \{0,1\}^3$, and forms a belief of L's scruples $\mu^{a_p,y,z} = Pr(\theta = 1|a_p, y, z)$. In equilibrium, A punishes the leader if $\mu^{a_p,y,z}$ falls below the exogenous threshold $\bar{\mu}$.

Finally, considering the leader's payoff: L enjoys a benefit normalized to 1 for policy success (and 0 for policy failure); he receives a benefit of $\beta > 0$ for being rewarded by A (and 0 if punished), and pays costs k_p and k_c^θ for each respective action. Altogether,

$$U_L(a) = y - a_p k_p - a_c k_c^\theta + r\beta \tag{3}$$

Scrupulous and unscrupulous leaders differ only in the direct cost they incur from taking an unscrupulous action, k_c^θ; we will assume k_c^1 is arbitrarily high,[13] and for shorthand we will denote $k_c = k_c^0$. We analyze Perfect Bayesian Equilibria of the model, under two technical assumptions specified in the appendix.[14]

[13] It suffices to assume that $k_c^1 > 1 + \beta$.

[14] These assumptions serve to: 1) impose an upper bound β so as to support equilibria in which scrupulous leaders take public action when it is feasible, and refrain when it is not (that is, playing $a_p = \omega_p$ for $\omega_p = 0, 1$); and 2) restrict

2.2 Analysis: Covert Action and the Cover-Up Mechanism

We focus our analysis on conditions that give rise to a strategic tension for the leader, which the tools of game theory are especially well-suited to help us reason through.[15] Specifically, we focus on the case in which a leader believes that the policy options that are consistent with international norms and laws are unlikely to work; but the leader is presented with an attractive but unscrupulous covert operation that has some chance of working. (Formally, this is the case of $\omega_p = 0$ and $\omega_c = 1$, which are privately observed by the leader.) In this situation, the leader faces a choice between advancing some foreign policy objective through unscrupulous means, or forgoing his preferred policy outcome in order to avoid international and domestic backlash.

How does the leader resolve this tension? Below, we will show that there are three general strategies the leader can employ. But first, we formally define our novel mechanism:

Definition 1 (Cover-Up) *Given that a leader has taken a feasible covert action ($a_c = 1$ when $\omega_c = 1$): we say that the leader takes a cover-up action when he also takes an infeasible public action ($a_p = 1$ when $\omega_p = 0$).*

The game's equilibrium exhibits distinct behavior depending on the values of the exogenous parameters. These results can be expressed most succinctly as a function of λ, the baseline risk of covert action exposure.

Proposition 1 *There are three classes of equilibria to the game. In all equilibria, the scrupulous leader always refrains from both covert action and cover-up action, but the scrupulous leader's behavior differs:*

- *In a Class I equilibrium, the unscrupulous leader never takes covert action.*
- *In a Class II equilibrium, the unscrupulous leader takes covert action with positive probability, but never takes cover-up action.*
- *In a Class III equilibrium, the unscrupulous leader takes both covert action and cover-up action with positive probability.*

attention to these intuitively sensible equilibria.

[15] Several results do not generate difficult strategic challenges to overcome. For example, if the leader is given a morally palatable public action that is likely to supply a policy success, he takes it. If the leader is presented with a covert action that is very unlikely to supply a policy success, he does not take it. We report these cases in the Appendix.

There exist thresholds λ', λ''*, where* $\lambda' \leq \lambda'' = \frac{\alpha_c - \alpha_0 - k_c}{\beta}$*, such that:*

- *When* $\lambda > \lambda''$*, only the Class I equilibrium exists, and it is unique.*
- *When* $\lambda \in (\lambda', \lambda'')$*, only the Class II equilibrium exists, and it is unique.*
- *When* $\lambda < \lambda'$*, only Class III equilibria exist.*

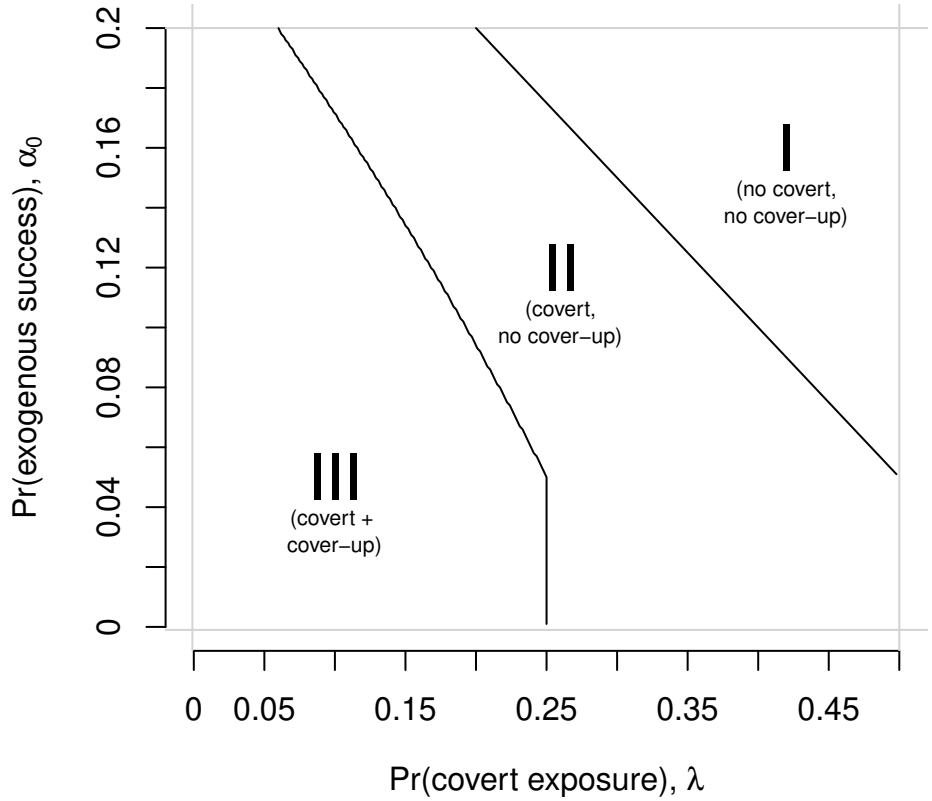
The basic logic of Class I and Class II equilibria are established in existing research. In Class I, the risk of exposure and the costs of covert action are not worth the potential policy gain. In this situation, the leader refrains from any covert or public action. In Class II, the risk of exposure is sufficiently low, and the policy gain sufficiently high, that the leader (at least sometimes) takes the gamble of a costly covert action.[16] The only risk the leader faces in this case is that of direct exposure; if the covert action succeeds, but the audience observes no direct evidence of it, they are willing to grant the leader the benefit of the doubt and infer that the success arose from exogenous factors.

Class III equilibria, in contrast, are entirely unstudied in existing work. In a Class III equilibrium, the leader (sometimes) takes covert action; and when he does, he (sometimes) also takes a public action, which does nothing to improve the odds of policy success and only serve to draw attention to the issue and raise the risk of exposure. To understand this counterintuitive behavior, it is useful to begin by examining the logic that supports the more standard behavior within the Class I and II equilibria, and then consider when that logic breaks down.

The conditions that differentiate the Class I equilibrium from the Class II equilibrium are fairly intuitive. If the direct risk of exposure is high ($\lambda > \lambda''$), then the leader is better off foreclosing that risk by refraining from covert action entirely. This condition is more likely to be satisfied as covert action becomes less effective ($\alpha_c \downarrow$)—or, alternatively, as the likelihood of success absent intervention increases ($\alpha_0 \uparrow$)—and as the leader's reputational concerns increase ($\beta \uparrow$). Conversely, when these conditions are not satisfied, the leader finds that the policy gains of covert action outweigh the reputational risks, and the equilibrium moves into the Class II region.

[16] Note that both Class II and Class III equilibria include conditions in which the leader plays a mixed covert action strategy, and plays a pure covert action strategy (i.e. always taking covert action), depending on parameter values. We group these different equilibria together under the same "class" for presentational clarity.

Figure 2: Equilibrium Regions



Note: Lines denote cutoff values for equilibria of Class I, II, and III, as defined in Proposition 1. Figure constructed with parameters: $\tau_p = 0.5, \tau_c = 0.8, \alpha_p = 0.8, \alpha_c = 0.4, k_p = 0.1, k_c = 0.1, \beta = 0.5, \pi = 0.5, \bar{\mu} = 0.4, \delta = 0.1$.

Eqm. Class	Covert Action	Overt Action	Supported if & only if	Summary of Leader's Reasoning
I	No	No	$\lambda > \lambda''$	The risk of exposure is high relative to the policy gain. Leader always avoids covert action.
II	Yes	No	$\lambda \in (\lambda', \lambda'')$	Policy gains are large. Leader is most concerned covert operators will be exposed during planning/execution phase. Leader avoids public action to avoid scrutiny over covert action.
III	Yes	Yes	$\lambda < \lambda'$	Policy gains are large. But people strongly suspect that policy success could only follow some form of intervention. Leader takes performative public action to reduce public suspicions after mission completes.

Within the Class II equilibrium, the unscrupulous leader employs covert action when covert action is feasible but public action is not—hoping that the covert action itself will not be exposed, and that if it achieves the desired outcome, the audience will attribute that outcome to exogenous

factors rather than to the leader's covert meddling. Two problems can arise, however. As λ gets large, the direct risk of exposure of the covert action increases, rendering the reputational gamble too risky to justify the potential policy gains. (This corresponds to the move from region II to region I in Figure 2.) But as λ gets small, the audience becomes increasingly suspicious of the leader's conduct, even in the absence of any direct evidence of covert action. (This corresponds to the move from region II to region III in Figure 2.)

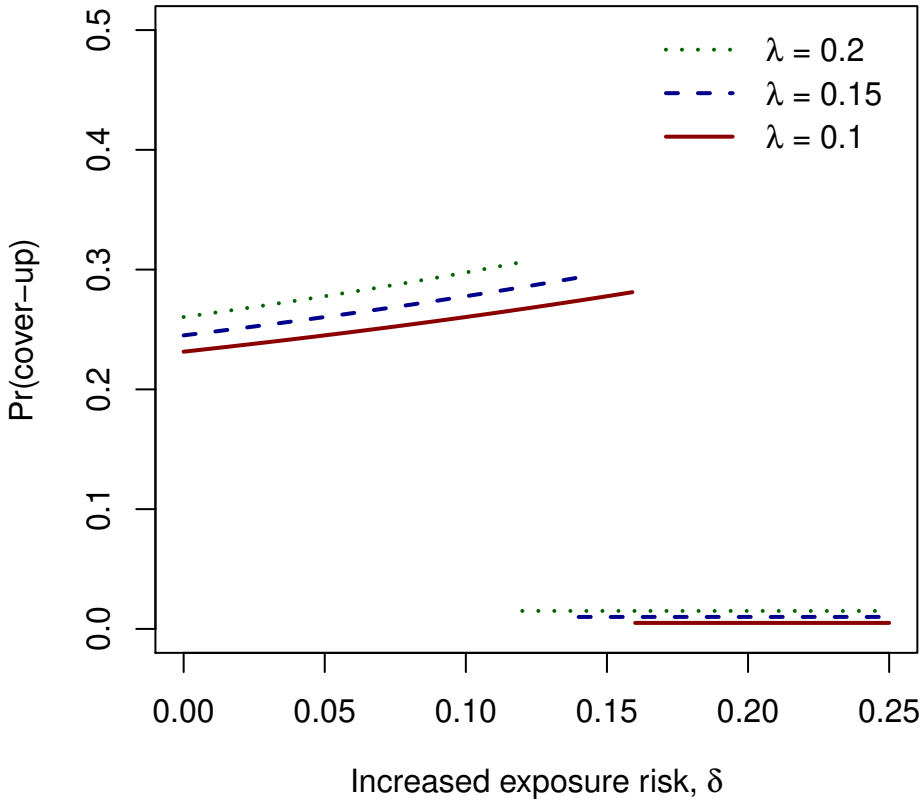
What explains the audience's increased suspicion as λ decreases? Upon observing the information set $(a_p = 0, y = 1, z = 0)$—that is, an outcome that the leader wanted, but with no public action taken by the leader toward that end, and no direct evidence of covert action—the audience infers that one of two things must have occurred: either the successful outcome came about due to random luck (which happens with probability α_0); or it came about as a result of the leader's unexposed covert action (which happens with probability $(1 - \lambda)\alpha_c$). As α_0 and/or λ decreases, the audience assigns greater probability to the more nefarious option.

If the audience's inference in this situation is too unfavorable, how will the leader respond? One option is to refrain from covert action entirely; but as discussed above, this is sub-optimal when the policy returns to covert action $(\alpha_c - \alpha_0)$ are too large to resist. Another option is to take a cover-up action. In addition to taking an effective covert action, the leader can also make a show of taking a public action that he (privately) knows to be ineffective. Then, upon observing the public action and a successful outcome, the audience attributes the success to the public action, rather than suspecting hidden foul play.

There are, however, important limitations on the leader's ability to skirt accountability through cover-up actions. First, and most fundamentally, the use of cover-up actions is partially self-defeating: when the audience expects the leader to use cover-up actions, they become increasingly skeptical of successful outcomes that are accompanied by public action, ultimately making covert action alone a more attractive option for reputational reasons. Put differently, the more likely a leader is to use a cover-up action, the less valuable it becomes to do so. This tension implies that any use of cover-up actions must be part of a mixed strategy, given the leader's strategic imperative to maintain unpredictability.

In addition, our model includes a technological assumption which serves to disincentivize cover-up behavior. Specifically, recall that the risk of direct exposure of covert action is λ in the absence

Figure 3: Increased exposure risk can increase cover-up incentives



Note: λ is the baseline risk of covert exposure; δ is the increased risk due to public action. Lines denote Pr(cover-up) within a Class III equilibrium (on lefthand side) and within the Class II equilibrium (on righthand side); discontinuities denote change across equilibria. Figure constructed with parameters: $\tau_p = 0.5, \tau_c = 0.8, \alpha_p = 0.8, \alpha_c = 0.4, \alpha_0 = 0.05, \kappa_p = 0.1, \kappa_c = 0.1, \beta = 0.5, \pi = 0.5, \bar{\mu} = 0.4$.

of public action, but $\lambda + \delta$ with public action (for some $\delta > 0$)—reflecting the intuition that public action draws public attention and media scrutiny, from within both the target state and the intervening state. If this increased exposure risk δ is too large, then the leader will refrain from taking any cover-up action. Yet we find that, even for moderate δ values, the leader nonetheless finds it beneficial to take an ineffective public action that raises the risk of direct exposure—a benefit that arises entirely due to the audience's strategic inference about the leader's behavior.

In fact, there exist conditions under which the leader's willingness to take a cover-up action is actually *increasing* in the direct exposure risk that accompanies it. This phenomenon is illustrated in Figure 3. Consider the effect of increasing δ from zero, starting on the lefthand side of the figure and moving rightward. Two things happen as an immediate consequence of δ increasing: first, by

definition, there is an increased risk that the leader's public action draws attention which causes his private action to be exposed, which disincentivizes cover-up action; but second, the audience's inference upon observing public action *without* covert revelation becomes more favorable, which incentivizes cover-up action. The left half of Figure 3 depicts an equilibrium in which the latter effect dominates.[17] This logic holds as δ increases up to a threshold, at which point the direct exposure risk is high enough to fully disincentive cover-up actions. The figure also depicts a similar logic with respect to the baseline risk of exposure λ: as λ increases, the leader increases his use of cover-up actions, up to the point that he abandons them entirely.

2.3 Empirical Predictions

As discussed above, a number of our model's implications—particularly those relating to the choice of whether or not to use covert action (i.e. the distinction between Class I and Class II equilibria)—have been developed in previous research. In addition, our model's focus on informational asymmetry and strategic inferences yields a number of novel empirical implications.

First, we should observe that decision-makers within Intervener states exhibit a concern for strategic inferences drawn by the audiences among whom they hope to maintain a scrupulous reputation. That is, when planning covert actions, decision-makers will not only consider the operational security and the risk of direct exposure of those actions. They will also consider how they are perceived by a skeptical audience—even in the best-case scenario that the operation succeeds with no direct exposure—and how they might be able to allay the audience's suspicion of their involvement.

Second, when the level of transparency (λ) around a foreign policy issue is low, or the odds of a favorable outcome absent intervention (α_0) are low, the Intervener will seek out cover stories for their covert interventions. The cover story should be a policy that is not clearly a violation of international laws and norms (or at least, a lesser violation than the covert action being pursued); it should not be too intrinsically costly (k_p low); and it should have a plausible nexus to the outcome being pursued through covert action. Third, when a covert action has succeeded in achieving the desired policy outcome, and no direct evidence of the operation has been revealed, the Intervener

[17] Note that this equilibrium is not unique; under these same conditions, there also exists an equilibrium in which Pr(cover-up) is constant in δ, up until the discontinuity depicted in the figure.

should make an effort to connect the outcome to the cover story in the mind of the audience.

We examine each of these implications in the case study that follows.

3 Operation PBSUCCESS

The 1950 presidential election was pivotal moment for Guatemalan democracy. Guatemala's first election came in 1944, following protests, and a revolutionary coup that brought down a long-standing and brutal dictator (Immerman, 1982, 34-35). The winner of the 1944 election, Juan José Arévalo, was the main revolutionary leader (Gleijeses, 2021, 36). The 1950 election was therefore the first time that Guatemalans had come to the polls where the main revolutionary leader was not on the ballot. Jacobo Arbenz won the election by a wide margin, [18] and assumed the presidency through an orderly transfer of power. From an institutional perspective, the 1950 election suggested that democracy was working in Guatemala (Fraser, p487). But it was not working for the United States. Answering the calls of the Guatemalan communist party, Arbenz implemented extensive land and agrarian reforms (Schlesinger and Kinzer, 1982, 53), which directly affected US commercial and political interests. US policymakers were also concerned by the many communists appointed to key government positions (Immerman, 1982, 108). Reports from CIA suggested Arbenz would provide the Soviet Union a major foothold in Latin America, through which they could spread communism.

In August, 1953, Eisenhower authorized the covert CIA operation PBSUCCESS. In its first phase, the CIA established bases in neighboring countries. They used these bases to train and arm 480 Guatemalans to launch a coup. The CIA also groomed a staunched anti-communist and former coup plotter, Castillo Armas, to lead the rebellion. But the real genius of the plan lay in the psychological operations (Cullather, 2006). The CIA was skeptical that a small group could overthrow the government. Thus, the covert plan included offensive psychological operations to convince loyalists that defense of Arbenz was futile and would lead to reprisals. This included a media blitz across Latin America, bribes to Guatemalan politicians to have them recognize the coup plotters as the rightful governments, and threats against those whom they could not buy (Schlesinger and Kinzer, 1982, 114). The paramilitary operations only commenced after months of

[18] He received 65% of the vote. The next closest candidate received 18%.

psychological operations shook confidence in Arbenz. PBSUCCESS is widely seen as a successful covert action. In the face of military incursions, Arbenz resigned on 27 June 1954. The CIA avoided direct evidence of their involvement. Broadly speaking, the US retained enough plausibly deniability to avoid backlash.

We use this case to lend support to our theory. Generally speaking, our main questions are: did the Eisenhower Administration worry about strategic inferences? If it did, how did it avoid them? But we have many other questions along the way. Most specifically, we derive our empirical approach from best-practices in using case material to illuminate equilibrium predictions from formal models. Following Bates (1998) we searched for a case where the initial conditions are consistent with the parameter ranges necessary to support our most interesting equilibrium (the cover story equilibrium). This led us to PBSUCCESS. Once selected we read through the decision-making logic of the case. Following Goemans and Spaniel (2016); Joseph, Poznansky, and Spaniel (2022) and others we read through decision-making processes in the case, paying particular attention to the choice nodes that we model. We developed case-specific hypotheses about what our theory expects to see. In evaluating those hypotheses, were were careful to consider alternative explanations from existing theoretical research into covert action.

3.1 Calibrating the parameters

We initially focused on cases in which Eisenhower faced a choice to oust popular governments that supported communism. This broad focus well fits many initial conditions of the cover story equilibrium. Eisenhower's overt options were constrained for fear of international backlash (Rabe, 1988, 166). The Liberal Order promoted sovereignty, self-determination and democracy. He realized it would be highly unscrupulous in the sense we mean it for the US to use military power to overturn democratically elected governments, or other governments with popular support (Poznansky, 2019, 86). Especially in Latin America, he worried that overt meddling would face unacceptable backlash. The Organization of American States was a fledgling institution, and populism was rising across Latin America. Eisenhower worried that brazen regime change would sour the opinions of Latin American elites and publics.

But Eisenhower's main foreign policy objective was to stop the spread of communism. Eisenhower believed that if he did not meddle, that communism would expand across the developing

world anyway (Schmitz, 1999, 179). To be clear, Eisenhower legitimately believed that sophisticated states if left to their own devices would prosper under democracy and support parties that wanted closer relations with the United States. But he also thought that developing states were unsophisticated, and as the result of Soviet meddling, were not left to their own devices. As a result, he was gravely concerned that absent forceful US intervention, communist parties would take hold in much of the developing world (Schmitz, 1999, 182).

Historians debate how calculating Eisenhower was. Early works questions how involved he was in foreign policy choices, his world view, and his competence (Divine, 1981). This work suggests key advisers, notably the Dulles Brothers (with Allen Dulles as CIA director, and John Foster Dulles as Secretary of State), played the major role. More recent work suggests Eisenhower was skillful and directly involved (see McAuliffe, 1981). The fact that this debate persisted helps us calibrate a parameter. After all, our cover story equilibrum assumes that the public and foreign audiences are uncertain about these facts. If it took historians decades to understand Eisenhower, it is plausible that audiences were uncertain about him in 1954.

Since this debate exists, we analyze documents surrounding the Administration's reasoning as a whole. But it is reassuring for our theory that Eisenhower's core advisers, the Dulles brothers, also took an instrumental view of the costs and benefits of covert regime change (Immerman, 1982, 22). While senior elites did not agree on everything, the fact that they score similarly on the dimension our theory cares about is useful for adjudicating the documents.

We focused on Guatemala specifically for two reasons. First, Eisenhower was in power at all stages of the decision making process.[19] Second, the case presents initial conditions that both illustrate it's importance and meet the conditions we believe our cover story equilibrium could arise. Notably, Guatemalla was especially high stakes because it represented the first major communist foothold in the Americas, putting the Soviets close to the US homeland.[20] Furthermore, the perceptions of costs for the overt and covert options match our parameter ranges. It was widely believed that US support for any military operations would lead to sever 'international consequences,' because Arbenz was democratically elected by a wide margin. By contrast, overt operations such as

[19]Eisenhower approved planning for the mission to oust Castro. But Kennedy approved the mission. Many features of Eisenhower's reasoning in the Cuba case fit our cover story equilibrium. But the US regime shift complicates the analysis. For an interesting overview, see CIA, Official History of the Bay of Pigs Operation, V II, pp12-14. Similarly, Truman conducted overt negotiations with Mossadeq in Iran before covert action was a possibility.

[20]The stakes for the US in Iran, Congo and elsewhere were likely not as large.

economic sanctions would lead to less backlash.[21]

3.2 Plausible deniability and Overt Action

Right from the outset, plausible deniability was essential. Whenever the mission was reviewed by the Administration, they reminded mission planners, "don't get caught."[22] Consistent with existing theoretical arguments (Joseph and Poznansky, 2018), this included diligent efforts to avoid direct evidence of US involvement. According to Immerman (1982, 133) (p133) "Planning took place with the utmost stealth. Only Eisenhower, the Dulles brothers, and a few other top-level members of the White House, State Department, and Central Intelligence Agency knew that an operation was even being considered, let alone were privy to its details."

But there is only so much careful planning can do. There was always a risk of direct exposure because CIA officers were stationed across Latin America to train and supply coup-plotters. In December 1953, the CIA even opened an operation center in Guatemala (Cullather, 2006, App. A). After the active phase of PBSUCCESS was given the "full green light" in April 1954, CIA officers remained in Guatemala and South America to facilitate psychological operations, bribe Guatemalan politicians and military officers, and otherwise monitor the plot (Cullather, 2006).

Given the intense focus on sustaining secrecy, we might expect that the Executive would divert public attention away from Guatemala as CIA officers were in the field so as to minimize the risk of direct exposure. But that is not what happened. In early 1954, the US Ambassador to Guatemala (Peurifoy) and others made inflammatory statements that the US would not tolerate a communist country between Florida and the Panama Canal. In March, at the Caracas Conference of the OAS, Eisenhower forced an anticommunist resolution first on the agenda designed to isolate Guatemala (Immerman, 1982, ch 19).

During the military phase of PBSUCCESS, when the CIA was most exposed, the Administration increased their overt policies. On May 15, a freighter carrying weapons that Arbenz had purchased from Czechoslovakia landed in Guatemala (Immerman (1982, 155); Schlesinger and Kinzer (1982, 147)). Arbenz had hoped to keep the shipment a secret, but the US discovered it the very next day (Cullather, 2006, 80).[23] Rather than minimize the episode, Eisenhower expressed public outrage.

[21] See Memorandum for Col J. C. King, PBSUCCESS 20th Jan 1954.

[22] https://history.state.gov/historicaldocuments/frus1952-54Guat/d116

[23] Obviously, the CIA did not plan this secret weapons shipment. They only exploited it when they found it. How-

He invoked the Monroe Doctrine which called for the exclusive influence of the United States in Latin America. He then announced a blockaded against all shipping into Guatemala (Cullather, 2006, 79).[24]

At the same time, the US convened an emergency meeting of the Organization of American States in which Dulles delivered an impassioned anti-Guatemalan speech. This was at Eisenhower's direction, who instructed his diplomats that "By every proper and effective means we should demonstrate to the courageous elements within Guatemala who are trying to purge their government of its communist elements that they have the sympathy and support of ... the U.S. ..." By proper, he meant public and short of calling for military intervention (Bowen, 1983). After months of delay, the Executive also authorized a Memorandum of Understanding with Honduras on military exchange, with the view of enhancing protection from neighboring communist states (i.e. Guatemala).

Why would Eisenhower shine a light on Guatemala when the CIA was most exposed? The known explanation is that mission planners wanted to maximize the chance Arbenz would step down by maximizing the psychological pressure on Arbenz and minimizing his capacity to resist the paramilitary operations. This incentivized Eisenhower to authorize all available policies, both overt (but short of military intervention) and covert (e.g. Cullather, 2006, p59).

We agree with this argument. On its own, it is not inconsistent with our theory. We account for the possibility that overt and covert policies can raise the chance of success independently. The important question for us is: does Eisenhower's desire to deploy all policy tools fully account for the extent of overt actions we observed? If the answer is yes, we would expect the Administration to only publicize overt policies when it confers operational advantage. We believe that two aspects of how Eisenhower publicized overt actions do not.

First, the Executive publicized events within the United States. In fact, DCI Dulles deliberately exaggerated the scope of the weapons shipment to prompt Congressional statements and press coverage (Cullather, 2006, p59). There were operational disadvantages to engaging the US public directly. One concern was that LINCOLN was commanding operations from an undisclosed location in Florida. One of the satellite offices had recently been closed due to potential surveil-

ever, if no pre-text arose, the CIA had planned to fabricate a phony Soviet arms Cache under operation WASHTUB (Cullather, 2006, 101).

[24] Consistent with our parameter ranges, these actions involved a small normative cost because Latin American states thought that South American states could buy arms from any supplier. Thus, the US naval blockade was unjustified. However, the backlash was relatively small (Friedman, 2010, 672).

lance concerns.[25] The more attention within the United States, the more media scrutiny would follow, raising the chance of exposure at this critical operational moment. Different still, Assistant Secretary of State Cabot had previously warned that if US 'public opinion should become too aroused and excited, there might be an embarrassing demands for action... [that were] altogether infeasible."[26]

Second, while PBSUCCESS relied partly on broadcasting anti-Arbenz messages across Guatemala, mission success did not rely on messages voiced from American foreign policy elites. In fact, there was concern that 'hard hitting speeches against Guatemala by personages in the United States Government could be counter-productive and would particularly alienate those non-Communists whose actions are influenced by nationalist emotions.'[27] Thus, it is not clear why Eisenhower would call on diplomatic staff to directly voice anti-Guatemalan positions when PBSUCCESS was operating local radio stations that could have voiced the same messages.

3.3 The cover story explanation.

We argue that the Administration was also concerned about strategic inferences. While they knew that additional attention would raise the risk of direct exposure, and even cause mistrust amongst Guatemalan nationalists, this risk was necessary to create a story that would allow the US to disclaim coup plotting in the years after the operation was complete.

If we are right, we will see three things. First, the Administration should directly acknowledge the problem of strategic inferences during the planning phase of PBSUCCESS. This is what we see. While planning the operation, the NSC explicitly acknowledged that even if no direct evidence of CIA involvement was revealed, "It must be recognized that any major effort to dislodge the Communist-controlled government of Guatemala will probably be credited to the United States, and possibly on CIA." As a result, "Covert accomplishment of the objectives of PBSUCCESS is therefore defined as meaning accomplishment with plausible denial of United States or CIA participation"[28] after the operation was concluded. Consistent with our theory, the NSC defined success in terms of overall perceptions of US involvement even absent direct evidence.

[25] I HAVE THIS DOCUMENT BUT I HAVE LOST IT!!!! I WILL FIND IT.
[26] Wisner's Memorandum for Chief, Western Hemisphere Division, 30 December 1953.
[27] Synthesis of [classified]'s remarks relevant to PBSUCCESS. April 22, 1954.
[28] https://history.state.gov/historicaldocuments/frus1951v02/d804

Second, we expect political staff who make public statements about Guatemala to point to overt actions to disclaim regime change operations. This is what we see. For example, Second Secretary of Embassy in Guatemala, Hill, recounted his conversation with an anti-Communist Guatemalan (whose name is still classified) as follows, "I told [Classified] that Ambassador Patterson had been quite correct in pointing out the US policy of non-intervention... but [Classified] was quite wrong in thinking that the US was not seriously concerned about the communist problem here... Assistant Secretary Cabot and others had made our concern with Communism in Guatemala abundantly clear in recent speeches; and we were now seeking means to combat Communism on a hemispheric basis through cooperation with other Latin American nations at the forthcoming Caracas Conference." Hill then explained, "In talking in this vein to [classified] it was my intention to give him the impression that the US had no concrete plan for intervention in the domestic affairs of Guatemala and continued its non-intervention policy..."[29]

Similarly, we expect that mission planners to acknowledge the problem of strategic inferences and that to the extent that they are severe, that cover stories can play a role to overcome them. The issue arose during the critical decision period of April 1954 when DCI Dulles was considering authorizing the final stages of PBSUCCESS. CIA Deputy Director for Plans, Wisner, thought it was an ideal time to authorize military operations. Arbenz was weakened by successful psychological operations, but a recent NIE suggested he was tightening controls. Thus, if CIA did not act now, they may never have the opportunity. However, Assistant Secretary for State Holland argued for delay because the risk of exposure was too great. Security Officers had found recording devices in the homes of CIA staff, the US Ambassador to Honduras (Willauer) had shared too much information about US operations with friendly Honduran officials, and Arbenz had stumbled across suggestive evidence of a US plot to fabricate a Soviet Weapons cache (Cullather, 2006, p42).

On April 21st, Wisner agreed with Holland that "documentary evidence may not be necessary to establish the intervention case against the United States... a strong circumstantial case could be as effective as actual evidentiary material."[30] But Wisner still wanted to push ahead. On the 24th he summarized the debate in a Position Paper for DCI Dulles. Again, he acknowledged that the question to advance PBSUCCESS to the military phase hinged on concerns about strategic

[29]Dispatch No. HGG-A-619. Memorandum of Conversation, From COS, Guatemala to Chief WHD. Jan, 11, 1954. For a second example, see Memorandum of Conversation from Hill to Charge d'Affaires, Jan 25, 1954.

[30]Minutes of Weekly PBSUCCESS Meeting. 21 April, 1954.

inferences. He assessed that "It is fair to assume that no irrefutable evidence tying the project to the U.S. Government is in the hands of the enemy." However, "there is not the slightest doubt that if the operation is carried through many Latin Americans will see in it the hand of the U.S."

For him, the question was to what extent will these audiences attribute the coup to the United States, and what could be done about it. He then laid out three positions. He characterized Holland's position as a substantial delay to develop the most comprehensive cover story. "What is specifically proposed is to begin with a strong official statement of the United States position toward the present regime in Guatemala, followed by an attempt to secure the support of the Caracas majority at an OAS meeting in September for the application against Guatemala of the sanctions envisaged in the Rio Pact. Essentially this would involve an economic and communications blockade of Guatemala by OAS members or at least by those members willing to support and join in the action."[31] In short, by acting through the OAS, imposing sanctions, and other actions, the US could credibly explain that they were not taking covert actions because they had a plan that they were openly pursuing. He also characterized a middle ground that involved a moderate delay with a moderately improved cover story involving "vigorous and coordinated program of official and overt action and covert operations."

Instead, Wisner argued in favor of moving ahead with minimal modifications to PBSUCCESS. However, his recommendation did not disclaim the value of a cover story. Rather, he assessed that, "The security of the project is as good as can be expected and fully in keeping with the estimates made and reported on numerous occasions starting with the beginning of the project." In fact, a month earlier, he had explicitly argued that "it might be a good idea to cry wolf several times before D-Day."[32] In June, Wisner's subordinates who were managing the operations from LINCOLN observed with disappointment that US Ambassadors in Honduras and Guatemala were not publicly voicing the US position. They thought it was "essential that for diplomatic battle the hole created by non-participation should be filled."[33]

Putting it altogether, we think this debate well supports our argument. As Wisner characterizes it, the level of concern about strategic inferences dictates the need for developing a more comprehensive cover story. He believed that enough counter-measures were in place, and therefore delay

[31] https://history.state.gov/historicaldocuments/frus1952-54Guat/d133
[32] PBSUCCESS Weekly Meeting, 30 March 1954.
[33] Memorandum from LINCOLN to RYBAT (CIA) Director, 21 June, 1954.

was unnecessary. But he also argued that those with greater concerns about strategic inferences would want more comprehensive cover stories to off-set those concerns.

Finally, we expect that the after the mission was complete, the US will publicize the overt actions it took as a means to cover up their covert action. This is what we see. An NSC report, later released to the press, read:

> The Organization of American States was used as a means of achieving our objectives in the case of communist intervention in Guatemala. After the arrival from Poland on May 15 in Guatemala of a substantial shipment of arms, the United States initiated consultations with all Laten American Governments, except Guatemala. Following these consultations, the Council of the Organization of American States voted almost unanimously... to convoke a Meeting of Ministers of Foreign Affairs. The Council of the OAS postponed the meeting sine die because the revolution in Guatemala overthrew the communist-controlled Government.
>
> The revolution in Guatemala caused the communist-controlled Arbenz Government to appeal to the UNSC and to the Inter-American Peace Committee of the OAS alleging aggression on the part of Honduras and Nicaragua, supported by other foreign nations. The United States took the position that the Organization of American States was ready, willing and competent to respond to the appeal. The Security Council voted (Soviet Union against), in effect, to leave the matter to the OAS. The Inter-American Peace Committee prepared to investigate, but before the Committee arrived in Guatemala, the new government of that country indicated that the controversy requiring the investigation had ceased to exist.[34]

As the quote shows, the story is that the US action was through the OAS, and therefore that they were not involved in the coup plotting.

What is more, it appears that analysts why studied US actions at the time refer to these actions to off-set suspicions. For example, 2 years after Arbenz was ousted, Taylor (1956) published a comprehensive 'Critique of United States Foreign Policy' in Guatemala surrounding Arbenz' removal. This critique included a full review of journalistic inquiries into US policies, South American and

[34] https://www.cia.gov/readingroom/docs/CIA-RDP80R01731R003000030008-2.pdf Operations Coordinating Board Feb 1955. Progress Report on NSC 5431/1 (Latin America)

academic policy studies into the US role. During this review, Taylor uncovers how a handful of pundits who conjectured that the CIA did support Armas covertly. "But it is difficult to find evidence which would clearly implicate [US ambassador to Guatemala] Peurifoy or other United States' representatives in the plotting which resulted in Castillo's invasion from Honduras." Consistent with our argument, his detailed review focuses, on US efforts to rouse anti-communist sentiment in the OAS as Eisenhower's policy intervention, and other diplomatic actions promoted by Secretary of State Dulles.

Putting it altogether, this evidence shows that the Executive is concerned about strategic inferences, that as part of the mission planning, the CIA conceptualized a diversionary public action so that they could retain plausible deniability in the face of these strategic inferences. Further, we see evidence that they refer back to these overt policies to divert attention away from their sponsorship in the years after the coup succeeds.

3.4 Three Clarifications

First, we want to clarify that the cover story mechanism was only one way that mission planners avoided strategic inferences. The CIA deliberately trained Guatemalan exiles to make the coup appear like a local conflict between Guatemalan political factions. The CIA also crafted the appearance of alternative foreign sponsors. Most notably, the CIA deliberately trained and armed the Guatemalan coup plotters in Nicaragua, Honduras and other countries that were hostile to Arbenz. Of course, training forces overseas raises the risk of direct exposure because the CIA cannot easily control the environment.[35] The CIA also armed the coup-plotters with weapons it purchased from the Dominican Republic to implicate them.[36] We view these methods as consistent with our overall theory. After all, each method raised the risk of direct exposure against the benefits from reducing strategic inferences. As the Operations Coordinating Board put it in a Memo designed to assess plausible deniability, the foreign training bases in Nicaragua "added support in cloaking the U.S. hand exists in the number of other countries which both have good reasons for wanting to

[35]There was a near miss in January 1954, when chatter from Nicaraguans privy to local operations prompted Guatemala to published a White Paper accusing 'the government of the North,' of supporting covert, anti-Guatemalan activities in Nicaragua. However, the chatter was unsubstantiated, and could have referred to Mexico. According to the CIA, 'Continued study of the aftereffects of the White Paper indicates that it somewhat reinforced suspicions among all those previously inclined to suspect the U.S. but was roundly disbelieved by the majority of anti-Communists in Central America."

[36]Memorandum for Goodbourne, Progress Report - PBSUCCESS. 11 May, 1954.

see the replacement of the Arbenz Government and have the means for backing a coup of the size planned."

Second, our analysis confirmed that some pundits speculated about US involvement shortly after Arbenz fell (eg Grant, 1955). Thus, Eisenhower did not completely escape strategic inferences. This is supportive of our mechanism. In our theory, public actions do not entirely prevent strategic inferences. Rather, they off-set suspicion enough so that the Intervener can avoid backlash. Consistent with our theory, some suspicion arose. But that suspicion was not wide-spread. We found no evidence that the US public, or Congress suspected Eisenhower's involvement. We also found no evidence that major foreign audiences within Guatemala, Latin America, or beyond seriously suspected US involvement. Rather, the suspicion was confined to a handful of academic publications and non-major newspapers.

A final concern is that Eisenhower engaged others at the OAS to offset backlash in the event that US actions became public. This would not be inconsistent with our argument if this objective followed along side the cover-story objective. However, it is notable that we found evidence of the cover story mechanism in NSC deliberations, and exchanges between Eisenhower and Dulles. We did not find any discussion of gaining consensus in the case that the covert action was exposed. It is also worth noting that this logic could not explain Eisenhower's choice to publicise the blockade or make inflammatory statements against Arbenz outside of the OAS meetings.

4 Conclusion

We argued that international and domestic audiences actively investigate world affairs. From their investigations, they try to infer from context and outcomes, not only direct evidence, the secret activities of powerful states. This insight illuminates a core tension states face if they hope to sustain plausible deniability for their secret policies. On the one hand, they must avoid getting caught as they plan and execute their operations. On the other, after they succeed, they need to explain how the world turned in their favor by chance, and not as the result of a repugnant action they took in secret. How can they simultaneously achieve these goals? When policy success is unlikely to follow absent intervention, a state's best chance for avoiding attribution of their secret policy is a cover story. The ideal cover story: (a) is more consistent with international laws

and norms than the secret policy the leader wants to deploy (e.g. public sanctions and covert assassinations), and (b) the public believes that it could, in theory, drive success. Cover stories are rational mechanisms to deny morally repugnant secret policies even if they invite additional scrutiny that raises the risk a covert mission is exposed, and are somewhat unpopular amongst international audiences. In fact, the more scrutiny they invite, the more effective they are at disclaiming secret policies ex-post.

Planning that surrounds Operation PBSUCCESS, the CIA mission to oust Arbenz of Guatemala, supports our theory. We find evidence that Eisenhower and Dulles worried about strategic inferences. They believed that many in the region would assume the US was responsible even if no evidence emerged of CIA involvement. They utilized many actions to avoid attribution. But an overlooked mechanism is a series of performative overt policies. Consistent with our theory, in the years after Arbenz resigns, the administration points to these open policies to explain what they were doing, and help discredit speculation that a nefarious secret plot went on.

Our theory holds direct implications for three kinds of political actors. First, it explains to CIA mission planners what they require from policymakers to retain plausible deniability if they intend to assassinate Iranian and North Korean scientists who are essential for those nuclear programs, or sow dissent among Chinese elites as Sino-American tensions flare. If they want these Targets to remain uncertain of US involvement, they need the State Department and military to take provocative public actions to create a plausible alternative story for what they were doing.

Second, it helps public accountability activists appreciate that building extensive monitoring capabilities may work against them in some cases. If the public widely believes that these organizations and the media can effectively scrutinize the government most of the time, then the public will infer that absence of evidence means no immoral policy took place. This, in turn, may make covert action more attractive.

Finally, it helps policymakers understand how to combat conspiratorial beliefs within the US, or sow these beliefs in rival states if they want to. If we are correct, then conspiracy theories could be a rational inference when implausible events occur that favor powerful states. Thus, psychological reprogramming and altering elite cues may not influence public beliefs. Rather, we need to convince the public that these seemingly rare events are plausible. If we can't we need to generate cover stories to help explain why we were not involved.

Our theory likely holds broader theoretical implications. There are many scenarios in American politics where our politicians can pursue unscrupulous policies and avoid media coverage. In these situations, they may utilize a similar cover story mechanisms. Indeed, this would provide one reason why states promote policies that we learn, in hindsight, are entirely performative.

References

Ashworth, S. (2012, 6). Electoral accountability: Recent theoretical and empirical work. *Annual Review of Political Science 15*, 183–201.

Ashworth, S. and E. Bueno de Mesquita (2014). Is voter competence good for voters?: Information, rationality, and democratic performance. *American Political Science Review 108*(3), 565–587.

Axelrod, R. and R. Iliev (2014, 1). Timing of cyber conflict. *Proceedings of the National Academy of Sciences 111*, 1298–1303.

Backus, M. and A. T. Little (2020, aug). I Don't Know. *American Political Science Review 114*(3), 724–743.

Baliga, S., E. Bueno de Mesquita, and A. Wolitzky (2020, 11). Deterrence with imperfect attribution. *American Political Science Review 114*, 1155–1178.

Bates, R. H. (1998). *Analytic narratives*. Princeton University Press.

Bowen, G. L. (1983, may). U.S. Foreign Policy toward Radical Change: Covert Operations in Guatemala, 1950-1954. *Latin American Perspectives 10*(1), 88–102.

Bull, H. (2002). *The anarchical society : a study of order in world politics*. Columbia University Press.

Canes-Wrone, B., M. C. Herron, and K. W. Shotts (2001, 7). Leadership and pandering: A theory of executive policymaking. *American Journal of Political Science 45*, 532.

Carnegie, A. (2021, 5). Secrecy in international relations and foreign policy. *Annual Review of Political Science 24*, 213–233.

Carnegie, A. and A. Carson (2018, may). The Spotlight's Harsh Glare: Rethinking Publicity and International Order. *International Organization 72*(03), 627–657.

Carson, A. (2015). Facing off and saving face : Covert intervention and escalation management in the korean war. *International Organization Forthcomin*.

Carson, A. (2018). *Secret wars : covert conflict in international politics*.

Coe, A. J. (2018, 10). Containing rogues: A theory of asymmetric arming. *The Journal of Politics 80*, 1197–1210.

Colgan, J. D. (2021). *Partial Hegemony: Oil Politics and International Order*. Cambridge University Press.

Cullather, N. (2006). *Secret History: The CIA's Classified Account of Its Operations in Guatemala, 1952-1954*. Stanford University Press.

Divine, R. A. (1981). *Eisenhower and the Cold War*. Oxford University Press.

Downes, A. B. and M. L. Lilley (2010, may). Overt Peace, Covert War?: Covert Intervention and the Democratic Peace. *Security Studies 19*(2), 266–306.

Downs, A. (1964). Inside Beuracracy. Technical report.

Fang, S. (2008). The informational role of international institutions and domestic politics. *American Journal of Political Science 52*(2), 304–321.

Farrell, H. and A. L. Newman (2021, 2). The janus face of the liberal international information order: When global institutions are self-undermining. *International Organization 75*, 333–358.

Finnemore, M. (2003, 11). Changing norms of humanitarian intervention.

Fox, J. and K. W. Shotts (2009, oct). Delegates or Trustees? A Theory of Political Accountability. *The Journal of Politics 71*(4), 1225–1237.

Friedman, M. P. (2010, dec). Fracas in Caracas: Latin American Diplomatic Resistance to United States Intervention in Guatemala in 1954. *Diplomacy & Statecraft 21*(4), 669–689.

Gentzkow, M. A. and J. M. Shapiro (2004, aug). Media, Education and Anti-Americanism in the Muslim World. *Journal of Economic Perspectives 18*(3), 117–133.

Gleason, J. (2017, sep). Showing Off: Promise and Peril in Unilateral Policymaking. *Quarterly Journal of Political Science 12*(2), 241–268.

Gleijeses, P. (2021). *Shattered Hope: The Guatemalan Revolution and the United States, 1944-1954*. Princeton University Press.

Goddard, S. E. (2018). *When right makes might : rising powers and world order*. Cornell University Press.

Goemans, H. and W. Spaniel (2016, jan). Multimethod Research:. *Security Studies 25*(1), 25–33.

Grant, D. (1955). Guatemala and the united states' foreign policy. *Journal of International Affairs 9*, 64–72.

Hakakian, R. (2011). *Assassins of the Turquoise Palace*. Grove Press.

Houghton, V. (2019). *Nuking the Moon: And Other Intelligence Schemes and Military Plots Left on the Drawing Board*. Penguin Books.

Huth, P. K., S. E. Croco, and B. J. Appel (2011, 5). Does International Law Promote the Peaceful Settlement of International Disputes? Evidence from the Study of Territorial Conflicts since 1945. *American Political Science Review 105*, 415–436.

Ikenberry, G. J. (2000). After victory : institutions, strategic restraint, and the rebuilding of order after major wars. pp. 293.

Immerman, R. (1982). *The CIA in Guatemala*. University of Texas Press.

Izzo, F. (2022, oct). Ideology for the Future. *American Political Science Review*, 1–16.

Joseph, M. F. and M. Poznansky (2018, 5). Media technology, covert action, and the politics of exposure. *Journal of Peace Research 55*, 320–335.

Joseph, M. F., M. Poznansky, and W. Spaniel (2022, apr). Shooting the Messenger: The Challenge of National Security Whistleblowing. *The Journal of Politics 84*(2), 846–860.

Kinzer, S. (2003). *All the Shah's Men*. John Wiley & Sons.

Krcmaric, D. (2019, 6). Information, secrecy, and civilian targeting. *International Studies Quarterly 63*, 322–333.

Lake, D. A., L. L. Martin, and T. Risse (2021, 3). Challenges to the liberal order: Reflections on international organization. *International Organization 75*, 225–257.

Levin, D. H. (2021). *Meddling in the ballot box : the causes and effects of partisan electoral interventions*.

Lindsay, J. R. (2013, 7). Stuxnet and the limits of cyber warfare. *Security Studies 22*, 365–404.

Malis, M. (2023). Foreign policy appointments. Working paper. http://mattmalis.github.io/research.

Maskin, E. and J. Tirole (2004). The politician and the judge: Accountability in government. *American Economic Review 94*(4), 1034–1054.

McAuliffe, M. S. (1981, dec). Eisenhower, the President. *The Journal of American History 68*(3), 625.

Miller, G. J. (2005, 6). The political evolution of principal-agent models. *Annual Review of Political Science 8*, 203–225.

Morse, J. C. and T. Pratt (2022, oct). Strategies of Contestation: International Law, Domestic Audiences, and Image Management. *The Journal of Politics 84*(4), 2080–2093.

Nyhan, B. and T. Zeitzoff (2018, 10). Conspiracy and misperception belief in the middle east and north africa. *The Journal of Politics 80*, 1400–1404.

Oliver, J. E. and T. J. Wood (2014, may). Conspiracy Theories and the Paranoid Style(s) of Mass Opinion. *American Journal of Political Science 58*(4), 952–966.

Owen, J. M. (1994, 23). How liberalism produces democratic peace. *International Security 19*, 87.

Poznansky, M. (2019, jan). Feigning Compliance: Covert Action and International Law. *International Studies Quarterly*.

Poznansky, M. (2020). *In the shadow of international law : secrecy and regime change in the postwar world*.

Poznansky, M. (2022, 6). Revisiting plausible deniability. *Journal of Strategic Studies 45*, 511–533.

Rabe, S. G. (1988). Eisenhower and latin america : the foreign policy of anticommunism. pp. 237.

Ramsay, K. W. (2004). Politics at the water's edge: Crisis bargaining and electoral competition. *Journal of Conflict Resolution 48*(4), 459–486.

Reiter, D. and A. C. Stam (2002). *Democracies at war*. Princeton University Press.

Schlesinger, S. C. and S. Kinzer (1982). *Bitter Fruit: The Untold Story of the American Coup in Guatemala*. Anchor books. Doubleday.

Schmitz, D. F. (1999). *Thank God They're on Our Side: The United States and Right-wing Dictatorships, 1921-65*. University of North Carolina Press.

Schram, P. (2022, 3). When capabilities backfire: How improved hassling capabilities produce worse outcomes. *The Journal of Politics*.

Schultz, K. A. (2005). The politics of risking peace: Do hawks or doves deliver the olive branch? *International Organization 59*(1), 1–38.

Smith, A. (1998). International crises and domestic politics. *American Political Science Review 92*(3), 623–638.

Spaniel, W. and M. Poznansky (2018, 7). Credible commitment in covert affairs. *American Journal of Political Science 62*, 668–681.

Stein, A. (2000). The justifying state: Why anarchy doesn't mean no excuses.

Taylor, P. B. (1956, 9). The guatemalan affair: A critique of united states foreign policy. *American Political Science Review 50*, 787–806.

Terman, R. and J. Byun (2022, 5). Punishment and politicization in the international human rights regime. *American Political Science Review 116*, 385–402.

Uscinski, J. E., C. Klofstad, and M. D. Atkinson (2016, may). What Drives Conspiratorial Beliefs? The Role of Informational Cues and Predispositions. *Political Research Quarterly 69*(1), 57–71.

Appendix

Table of Contents

A Formal Appendix

Notation:

- Let $\omega = (\omega_c, \omega_p)$; let $\sigma_j(\omega) = Pr(a_j = 1|\omega)$ for $j = c, p$; and let $\sigma(\omega) = (\sigma_c(\omega), \sigma_p(\omega)) = (\sigma_c(\omega_c, \omega_p), \sigma_p(\omega_c, \omega_p))$.
- In the game sequence, L chooses actions $a = (a_c, a_p) \in \{0, 1\}^2$ simultaneously. In characterizing L's strategy, it will at times be useful to denote his public action strategy as being conditional on his covert action, $\sigma_p(\omega; a_c) = Pr(a_p = 1|\omega, a_c)$.

Assumption 1 (Parameter restriction) *Assume:*

(i) $\kappa_c^{\theta=1} > 1 + \beta$;

(ii) $0 < \alpha_0 < \alpha_c - k_c^{\theta=0} < \alpha_p - k_p$

(iii) $\alpha_c < \alpha_p < 1$

(iv) $\beta < \min\left\{\frac{\kappa_p}{\alpha_0}, \frac{\alpha_p - \alpha_0 - k_p}{1-\alpha_0}\right\}$; *and*

(v) $0 < \bar{\mu} < \pi$.

As we will see further below, these assumptions imply, respectively:

(i) scrupulous leaders never take covert action;

(ii) absent reputational concerns, the unscrupulous prefers feasible public action over feasible covert action, and prefers feasible covert action over no action at all;

(iii) the leader has better information about the effectiveness of public action as compared to covert action, though neither can guarantee a policy success;

(iv) scrupulous leaders are willing to take public action when it is feasible, and willing to refrain from public action when it is not; and

(v) the leader will not be punished given the audience's unrevised prior belief of his scruples.

Assumption 2 (Markovian strategies) *We say that a strategy is Markovian if, whenever the leader receives the same payoffs for each possible action in two different information sets $\zeta = (\omega, \theta)$, he plays the same strategy in both information sets: that is, a Markovian strategy satisfies*

$$E[U_L(a)|\zeta] = E[U_L(a)|\zeta' \neq \zeta] \, \forall a \implies \sigma(\zeta) = \sigma(\zeta')$$

Restrict attention to equilibria in Markovian strategies.

Assumption 3 (Equilibrium restriction) *Restrict attention to equilibria characterized by the following behavior, when available:*

- *The leader always takes public action when feasible,* $\sigma_p(\omega_c, 1) = 1$.
- *When neither action is feasible, the leader takes neither,* $\sigma(0,0) = (0,0)$.

Lemma 1 *Under Assumption 1:*

- *The scrupulous leader never takes covert action.*
- *An equilibrium satisfying the conditions of Assumption 3 can always be supported.*

In any equilibrium satisfying Assumptions 1, 2, and 3:

- *The leader never takes covert action when covert action is infeasible,* $\sigma_c(0, \omega_p) = 0$*;*
- *The leader never takes covert action when public action is feasible,* $\sigma_c(\omega_c, 1) = 0$.
- *The leader only takes infeasible public action when he is simultaneously taking feasible covert action,* $\sigma_p(1,0) = a_c\widehat{\sigma}_p$.

Proof of Lemma 1: The lemma states that under Assumptions 1, 2, and 3, the following strategies must be played in equilibrium:

$\theta = 1$

ω	$\sigma(\omega)$
$(0,0)$	$(0,0)$
$(0,1)$	$(0,1)$
$(1,0)$	$(0,0)$
$(1,1)$	$(0,1)$

$\theta = 0$

ω	$\sigma(\omega)$
$(0,0)$	$(0,0)$
$(0,1)$	$(0,1)$
$(1,0)$	$(\widehat{\sigma}_c, a_c\widehat{\sigma}_p)$
$(1,1)$	$(0,1)$

with $\widehat{\sigma}_p \in [0,1]$ and $\widehat{\sigma}_c \in [0,1]$ left unspecified by the lemma.

First observe the following:

- The scrupulous and unscrupulous leader's best-responses are identical in the $(0, \omega_p)$ states.
- The scrupulous leader's strategy of $\sigma_c(1, \omega_p) = 0$ follows directly from the assumption that $\kappa_c^{\theta=1} > 1 + \beta$, meaning that the direct costs of covert action outweigh the highest possible reputational and policy gains from covert action.

Next, consider the audience's beliefs in each information set, given the strategies specified by the

lemma, using the notation $\mu^{a_p,y,z} = Pr(\theta = 1|a_p, y, z)$:

$$\begin{aligned}
\mu^{a_p,y,1} &= 0 \quad \forall a_p, y \\
\mu^{000} &= \frac{\pi}{\pi + (1-\pi)\left\{1 - \tau_c\widehat{\sigma}_c\left[1 - (1-\widehat{\sigma}_p)\left(\frac{1-\alpha_c}{1-\alpha_0}\right)(1-\lambda)\right]\right\}} \geq \pi \\
\mu^{010} &= \frac{\pi}{\pi + (1-\pi)\left\{1 + \tau_c\widehat{\sigma}_c\left[(1-\widehat{\sigma}_p)\frac{\alpha_c}{\alpha_0}(1-\lambda) - 1\right]\right\}} \leq \mu^{000} \\
\mu^{1y0} &= \frac{\pi}{\pi + (1-\pi)\left\{1 + \left(\frac{1-\tau_p}{\tau_p}\right)\tau_c\widehat{\sigma}_c\widehat{\sigma}_p\left(\frac{\hat{y}_c}{\hat{y}_p}\right)(1-\lambda-\delta)\right\}}
\end{aligned} \tag{4}$$

where $\hat{y}_j = \begin{cases} \alpha_j, & y = 1 \\ 1 - \alpha_j, & y = 0 \end{cases}$ for $j = p, c$. Observe that $\mu^{110} \leq \mu^{100}$.

We can denote the audience's reward/punishment strategy as $\hat{r}^{a_p,y,z} = Pr(r = 1|a_p, y, z)$.

From (2), we can see that $E[U_A(r = 1)|a_p, y, z] = \mu^{a_p,y,z}$, and $E[U_A(r = 0)|a_p, y, z] = \bar{\mu}$. Because $0 < \bar{\mu} < \pi$, it follows that in any equilibrium, $\hat{r}^{0,0,0} = 1$ and $\hat{r}^{a_p,y,1} = 0$.

Remark For notational convenience, let $q = \hat{r}^{010}$, $s = \hat{r}^{110}$, and $t = \hat{r}^{100}$.

Note that, as per Corollary 1 below, in any equilibrium satisfying Assumption 3 we have $s \geq q$ (that is, $\hat{r}^{110} \geq \hat{r}^{010}$).

Left to prove are the following incentive-compatibility conditions:

(i) $\sigma(0, \omega_p) = (0, \omega_p)$;

(ii) $\sigma(1, 1) = (0, 1)$; and

(iii) $\sigma_p(1, 0) > 0 \implies a_c = 1$.

(i) In either state $(0,\omega_p)$, $a_c = 1$ is strictly dominated by $a_c = 0$: it is directly costly, it does nothing to improve the odds of success, and it lowers the expected reputational payoff. So we will consider the choice of a_p in each state $(0, \omega_p)$.

For both leader types, we have

$$\begin{aligned}
E[U_L(a = (0, 0))|\omega = (0, 0)] &= \alpha_0 + \beta\left[\alpha_0\hat{r}^{010} + (1-\alpha_0)\hat{r}^{000}\right] \\
E[U_L(a = (0, 1))|\omega = (0, 0)] &= \alpha_0 - k_p + \beta\left[\alpha_0\hat{r}^{110} + (1-\alpha_0)\hat{r}^{100}\right]
\end{aligned}$$

The strategy $\sigma(0, 0) = (0, 0)$ is incentive-compatible if

$$\begin{aligned}
-k_p + \beta\left[\alpha_0\hat{r}^{110} + (1-\alpha_0)\hat{r}^{100}\right] &\leq \beta\left[\alpha_0\hat{r}^{010} + (1-\alpha_0)\right] \\
\alpha_0\left[\hat{r}^{110} - \hat{r}^{010}\right] - (1-\alpha_0)[1 - \hat{r}^{100}] &\leq \frac{k_p}{\beta}
\end{aligned}$$

The lefthand side of the inequality is $\leq \alpha_0$, so the inequality is satisfied for $\beta \leq \frac{k_p}{\alpha_0}$, as per Assumption 1.

In the state $\omega = (0,1)$, for both leader types, we have

$$E[U_L(a=(0,0))|\omega=(0,1)] = \alpha_0 + \beta\left[\alpha_0\hat{r}^{010} + (1-\alpha_0)\hat{r}^{000}\right]$$
$$E[U_L(a=(0,1))|\omega=(0,1)] = \alpha_p - k_p + \beta\left[\alpha_p\hat{r}^{110} + (1-\alpha_p)\hat{r}^{100}\right]$$

The strategy $\sigma(0,1) = (0,1)$ is incentive-compatible if

$$\alpha_0 + \beta\left[\alpha_0\hat{r}^{010} + (1-\alpha_0)\hat{r}^{000}\right] \leq \alpha_p - k_p + \beta\left[\alpha_p\hat{r}^{110} + (1-\alpha_p)\hat{r}^{100}\right]$$
$$\alpha_0\hat{r}^{010} + (1-\alpha_0) - \alpha_p\hat{r}^{110} - (1-\alpha_p)\hat{r}^{100} \leq \frac{\alpha_p - \alpha_0 - k_p}{\beta}$$

The lefthand side of the inequality is $\leq (1-\alpha_0)$, so the inequality is satisfied for $\beta \leq \frac{\alpha_p-\alpha_0-k_p}{1-\alpha_0}$.

(ii) Given that the scrupulous leader never takes covert action, it follows that in equilibrium, $\sigma(0,\omega_p) = (0,\omega_p) \implies \sigma(1,1) = (0,1)$ for the scrupulous leader.

For the unscrupulous leader, the assumption that covert action does not improve the odds of policy success when feasible public action is being taken[37] (that is, when $a_p\omega_p = 1$) implies that in state (1,1), he will choose one of the actions but not both. Thus in order to show that $\sigma_p(1,1) = 1$ can be supported in equilibrium, we want to show that $E[U_L(1,0)|\omega=(1,1)] \leq E[U_L(0,1)|\omega=(1,1)]$:

$$\alpha_c - k_c + (1-\lambda)\beta(\alpha_c q + (1-\alpha_c)) \leq \alpha_p - k_p + \beta(\alpha_p s + (1-\alpha_p)t)$$
$$(\alpha_c - k_c) - (\alpha_p - k_p) < \beta\left[\alpha_p s - \alpha_c q(1-\lambda) + (1-\alpha_p)t - (1-\alpha_c)(1-\lambda)\right]$$

Note that LHS of the second line is < 0, by Assumption 1.

There are three cases to consider. (1) If $\widehat{\sigma}_p = 0$, then $s = t = 1$, and RHS > 0. (2) If $\widehat{\sigma}_c = 1$ and $\widehat{\sigma}_p > 0$, then an equilibrium exists where $s = t = 1$, so again RHS > 0. (3) If $\widehat{\sigma}_c < 1$ and $\widehat{\sigma}_p > 0$, then that means

$$\begin{aligned} E[U_L(1,0)|(1,0)] &\leq E[U_L(0,0)|(1,0)] \\ &= E[U_L(0,0)|(0,1)] \\ &\leq E[U_L(0,1)|(0,1)] \\ &= E[U_L(0,1)|(1,1)] \end{aligned}$$

The first line follows from the fact that $\widehat{\sigma}_c < 1$, and the third line from point (i) above.

(iii) By Assumption 3, the leader's equilibrium strategy satisfies $\sigma(0,0) = (0,0)$. This means that $E[U_L(0,0)|\omega=(0,0)] \geq E[U_L(a)|\omega=(0,0)]$. Suppose that the leader has selected $a_c = 0$. Then $E[U_L(0,0)|\omega=(0,0)] = E[U_L(0,0)|\omega=(1,0)]$, and by Assumption 2, his strategy must satisfy $\sigma(1,0) = (0,0)$. By contraposition, $\sigma_p(1,0) > 0 \implies a_c = 1$.

■

Define $\Delta_p(\hat{r})$ as the leader's payoff from playing $a_p = 1$ over $a_p = 0$ given $\omega = (1,0)$ and $a_c = 1$,

[37] This assumption is a technical simplification, but our substantive results do not depend on it.

and given the audience's strategy $\hat{r} = \left(\hat{r}^{010}, \hat{r}^{110}, \hat{r}^{100}\right)$. That is:

$$\begin{aligned}\Delta_p &= E[U_L(a_p = 1)|\omega = (1,0), a_c = 1] - E[U_L(a_p = 0)|\omega = (1,0), a_c = 1] \\ &= -k_p + \beta\left\{\alpha_c\left[(1-\lambda-\delta)\hat{r}^{110} - (1-\lambda)\hat{r}^{010}\right] + (1-\alpha_c)\left[(1-\lambda-\delta)\hat{r}^{100} - (1-\lambda)\right]\right\} \qquad (5)\end{aligned}$$

Lemma 2 *In equilibrium,* $\widehat{\sigma}_p < 1$.

Proof: Suppose $\widehat{\sigma}_p = 1$. Then $\mu^{010} > \pi$, so $\hat{r}^{010} = 1$. But $\Delta_p\left(1, \hat{r}^{110}, \hat{r}^{100}\right) < 0$, meaning that L has a profitable deviation to $a_p = 0$ given $\omega = (1,0), a_c = 1$, contradicting $\widehat{\sigma}_p > 0$. So it must be that $\widehat{\sigma}_p < 1$. ■.

Corollary 1 *If* $\widehat{\sigma}_p = 0$, *then* $\hat{r}^{110} = \hat{r}^{100} = 1$. *If* $\widehat{\sigma}_p > 0$, *then* $\hat{r}^{110} \geq \hat{r}^{100}$ *and* $\hat{r}^{110} > \hat{r}^{010}$.

Proof: If $\widehat{\sigma}_p = 0$, then from (4) we see that $\mu^{110} = \mu^{100} = \pi > \bar{\mu}$. If $\widehat{\sigma}_p > 0$, then (given Lemma 2) this means that $\Delta_p = 0$. The quantity in the second set of square brackets in (5) is negative, so in order for $\Delta_p = 0$, the quantity in the first set of square brackets must be strictly positive, which requires $\hat{r}^{110} > \hat{r}^{010}$. And if $\widehat{\sigma}_p > 0$, then by visual inspection of (4) we see that $\mu^{110} > \mu^{100}$, meaning that either $\hat{r}^{110} = \hat{r}^{100}$ and both are 0 or 1, or $\hat{r}^{110} > \hat{r}^{100}$. ■

Lemma 3 *In equilibrium, if* $\widehat{\sigma}_c \to 0$, *then* $\widehat{\sigma}_p = 0$.

Proof: $\widehat{\sigma}_c \to 0 \implies \mu^{010} > \bar{\mu} \implies \hat{r}^{010} = 1 \implies \Delta_p < 0 \implies \widehat{\sigma}_p = 0$. ■.

We showed above that $\widehat{\sigma}_p < 1$. When considering the choice between $a_c = 0$ and $a_c = 1$, this means that the best payoff following from $a_c = 1$ can be achieved from $a_p = 0$.

So define Δ_c as:

$$\begin{aligned}\Delta_c(q) &= E[U_L(a_c = 1)|\omega = (1,0)] - E[U_L(a_c = 0)|\omega = (1,0)] \\ &= \alpha_c - \alpha_0 - k_c - \beta\left\{\lambda + (1-q)\left[(1-\lambda)\alpha_c - \alpha_0\right]\right\} \qquad (6)\end{aligned}$$

Lemma 4 *L's best-response covert action strategy in state* $\omega = (1,0)$, *and given audience strategy* $\hat{r}$, *is*

$$\widehat{\sigma}_c = \begin{cases} 0, & \Delta_c(q) < 0 \\ 1, & \Delta_c(q) > 0 \\ \widehat{\sigma}_c \in [0,1], & \Delta_c(q) = 0 \end{cases}$$

where $\Delta_c(q)$ *is given by* (6).

- *If* $\lambda \geq \bar{\lambda} := \frac{\alpha_c - \alpha_0}{\alpha_c}$ *then* $\mu^{010} \geq \pi$, *so* $q = 1$.
- *If* $\lambda < \bar{\lambda}$, *then* $\Delta_c(q)$ *is increasing in* q.

Proposition 2 *The leader takes covert action with positive probability,* $\widehat{\sigma}_c > 0$, *if* $\Delta_c(1) > 0$, *and only if* $\Delta_c(1) \geq 0$, *where* $\Delta_c(1) = \alpha_c - \alpha_0 - k_c - \beta\lambda$.

Proof of Proposition 2:

First, to prove the "only if", $\widehat{\sigma}_c > 0 \implies \Delta_c(1) \geq 0$:

- Suppose $\Delta_c(1) < 0$. Either (i) $\lambda < \bar{\lambda}$, in which case $\Delta_c(q)$ is increasing in q; so $\Delta_c(1) < 0 \implies \Delta_c(q) < 0 \,\forall\, q$, which $\implies \widehat{\sigma}_c = 0$; or (ii) $\lambda \geq \bar{\lambda}$, in which case $\mu^{010} > \bar{\mu} \,\forall\, \widehat{\sigma}_c, \widehat{\sigma}_p$, so $q = 1$, and $\Delta_c(q = 1) < 0 \implies \widehat{\sigma}_c = 0$. By contraposition, $\widehat{\sigma}_c > 0 \implies \Delta_c(1) \geq 0$.

Second, to prove the "if", $\widehat{\sigma}_c > 0 \impliedby \Delta_c(1) > 0$:

- Suppose $\Delta_c(1) > 0$. If $\widehat{\sigma}_c = 0$, then $\mu^{010} > \bar{\mu}$ so $q = 1$, and given $\Delta_c(1) > 0$, L has an incentive to deviate to $a_c = 1$, contradicting $\widehat{\sigma}_c = 0$. So $\Delta_c(1) > 0 \implies \widehat{\sigma}_c > 0$.

■

Corollary 2 *The condition that $\Delta_c(1) > 0$, which implies $\widehat{\sigma}_c > 0$, is satisfied for:*

- *high covert action effectiveness, α_c;*
- *low chance of random success, α_0;*
- *low transparency, λ;*
- *low reputational concerns, β*

Lemma 5

$$\textit{Let } \tilde{q} := \begin{cases} \frac{-\Delta_c(0)}{\Delta_c(1)-\Delta_c(0)}, & \lambda \leq \widehat{\lambda} \\ 1, & \lambda > \widehat{\lambda} \end{cases} \textit{, where } \widehat{\lambda} := 1 - \frac{\alpha_0(\psi - 1 + \tau_c)}{\alpha_c \tau_c}, \textit{ and } \psi := \frac{\pi(1-\bar{\mu})}{(1-\pi)\bar{\mu}}$$

Let $\widehat{q} = \max\{0, \min\{\tilde{q}, 1\}\}$.

If $\lambda < \bar{\lambda} = \frac{\alpha_c - \alpha_0}{\alpha_c}$, then:

- *$\tilde{q}$ is the unique solution to $\Delta_c(q) = 0$.*
- *$\Delta_c(1) > \Delta_c(0)$.*
- *$\Delta_c(1) \leq 0 \implies \tilde{q} \geq 1$.*

Proof of Lemma 5: μ^{010} is increasing in λ, with the lowest value occurring when $\widehat{\sigma}_p = 0$ and $\widehat{\sigma}_c = 1$. In this case, $\mu^{010} = \bar{\mu}$ rearranges to $\lambda = \widehat{\lambda}$, so any λ larger than this ensures $q = \hat{r}^{010} = 1$. The remaining points of the lemma follow from simple algebraic rearrangement of (6). ■

Proposition 3 *The leader uses a cover story with positive probability, $\widehat{\sigma}_c \widehat{\sigma}_p > 0$, if $\Delta_p(q = \hat{q}, s = t = 1) > 0$, and only if $\Delta_p(q = \hat{q}, s = t = 1) \geq 0$.*

Proof of Proposition 3:

First, to prove the "only if", $\widehat{\sigma}_c \widehat{\sigma}_p > 0 \implies \Delta_p(q = \hat{q}, s = t = 1) \geq 0$:

- Suppose $\Delta_p(q=\hat{q}, s=t=1) < 0$. Then we know that $\Delta_p < 0$ for all $q \geq \hat{q}$ and all s, t; thus $q \geq \hat{q} \implies \widehat{\sigma}_p = 0$. If $q < \hat{q}$ then $\Delta_c(q) < 0$, to which L's best response is $\widehat{\sigma}_c = 0$, which implies $q = 1$, contradicting $q < \hat{q}$. Thus $\Delta_p(q=\hat{q}, s=t=1) < 0 \implies \widehat{\sigma}_p = 0 \implies \widehat{\sigma}_c\widehat{\sigma}_p = 0$, and by contraposition, $\widehat{\sigma}_c\widehat{\sigma}_p > 0 \implies \Delta_p(q=\hat{q}, s=t=1) \geq 0$.

Second, to prove the "if", $\widehat{\sigma}_c\widehat{\sigma}_p > 0 \impliedby \Delta_p(q=\hat{q}, s=t=1) > 0$:

- (i). Suppose $\widehat{\sigma}_c = 0$ and show $\Delta_p(q=\hat{q}, s=t=1) \leq 0$.
  $\widehat{\sigma}_c = 0 \implies s = t = 1$. From Proposition 2, $\widehat{\sigma}_c = 0 \implies \Delta_c(1) \leq 0$. If $\lambda \geq \bar{\lambda}$ then $\tilde{q} = 1$; otherwise $\Delta_c(0) < \Delta_c(1)/$ in either case we have $\tilde{q} \geq 1$, which $\implies \widehat{q} = 1$. So $\Delta_p(q=\hat{q}, s, t) = \Delta_p(q=1, s=t=1) < 0$.
- (ii). Suppose $\widehat{\sigma}_c > 0$ and $\widehat{\sigma}_p = 0$, and show $\Delta_p(q=\hat{q}, s=t=1) \leq 0$.
  - If $\hat{q} = 1$, then $\Delta_p(\hat{q}, s, t) < 0$.
  - $\widehat{\sigma}_p = 0 \implies s = t = 1$.
  - From (5) we know that $\widehat{\sigma}_p = 0 \implies \Delta_p(q, s=t=1) \leq 0$. Left to prove is that $q = \hat{q}$.
  - If $\hat{q} < 1$, then either:
    * $\widehat{\sigma}_c = 1$, which (along with $\widehat{\sigma}_p = 0$ and $\lambda < \bar{\lambda}$) $\implies \mu^{010} < \bar{\mu} \implies q = 0$, and thus $\Delta_c(0) \geq 0$, meaning $\tilde{q} \leq 0 = \hat{q} = q$. Or:
    * $\widehat{\sigma}_c \in (0,1)$, meaning $\Delta_c(q) = 0$, so $q = \tilde{q} = \hat{q}$.

■

Proof of Proposition 1: From Lemma 1, Lemma 2, and Lemma 3, it follows that equilibria fall into one of three categories:

- Class I: $\widehat{\sigma}_c = 0, \widehat{\sigma}_p = 0$
- Class II: $\widehat{\sigma}_c > 0, \widehat{\sigma}_p = 0$
- Class III: $\widehat{\sigma}_c > 0, \widehat{\sigma}_p > 0$

From Proposition 2, we have that Class I equilibria are supported if and only if $\Delta_c(1) < 0$; and from Proposition 3, we have that Class III equilibria are supported if and only if $\Delta_p(s=1, t=1, q=\hat{q}) > 0$. It follows that Class II equilibria are supported if and only if both $\Delta_c(1) > 0$ and $\Delta_p(s=1, t=1, q=\hat{q}) < 0$.

First, observe that the condition that $\Delta_c(1) < 0$ rearranges to $\lambda > \frac{\alpha_c - \alpha_0 - k_c}{\beta} =: \lambda''$. Thus $\lambda > \lambda''$ implies that only the Class I equilibrium is supported.

Next, observe that

$$\Delta_p(s=1, t=1, \hat{q}) = -k_p + \beta\left[\alpha_c(1-\lambda)(1-\hat{q}) - \delta\right]$$

This can be evaluated by three cases of $\hat{q}$:

- If $\hat{q} = 1$, then $\Delta_p(1, 1, \hat{q}) < 0 \,\forall\, \lambda$, meaning $\lambda' = 0$.
- If $\hat{q} = 0$, then by visual inspection, $\Delta_p(1, 1, \hat{q})$ is decreasing in λ, and $\Delta_p(1, 1, \hat{q}) > 0$ rearranges to $\lambda < 1 - \left(\frac{k_p}{\beta} + \delta\right)/\alpha_c =: \lambda'$.
- If $\hat{q} \in (0,1)$, then $\hat{q} = \tilde{q}$, the unique solution to $\Delta_c(q) = 0$, as per Lemma 5.

In the third case, differentiating with respect to λ gives

$$\frac{d\Delta_p(1,1,\tilde{q})}{d\lambda} = \beta\left[-(1-\tilde{q}) - \frac{d\tilde{q}}{d\lambda}(1-\lambda)\right]$$

and $\frac{d\tilde{q}}{d\lambda}$ can be evaluated by the implicit function theorem as

$$\frac{d\tilde{q}}{d\lambda} = \frac{-\partial\Delta_c(\tilde{q})/\partial\lambda}{\partial\Delta_c(\tilde{q})/\partial\tilde{q}} = \frac{1-(1-\tilde{q})\alpha_c}{(1-\lambda)\alpha_c - \alpha_0} > 0$$

Thus $\Delta_p(1,1,\tilde{q})$ is decreasing in λ, and there exists a unique λ' such that $\Delta_p(1,1,\tilde{q}) > 0$ if and only if $\lambda < \lambda'$.

This exhausts all cases.

■