

Cover Stories

Michael F. Joseph
UCSD
mfjoseph@ucsd.edu

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

April 16, 2024

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 activities. We contribute by highlighting the challenges posed by *circumstantial* evidence: the inferences that audiences might draw from observing outcomes consistent with the intervener's interests and capabilities, but without direct evidence of their involvement. Through a formal model, we uncover a novel mechanism whereby interveners enhance their plausible deniability by openly promoting public coercive actions and, even, inviting some scrutiny that raises the risk of exposure. We find evidence of the use of such a "cover story" tactic in an in-depth case study of Operation PB-SUCCESS, the CIA operation that overthrew Guatemalan president Jacobo Arbenz in 1954. The theory advances our understanding of how leaders evade accountability for norm-violating behaviors across a wide range of foreign and domestic policy domains.

In March 1960, the CIA began organizing Cuban exiles to oust Fidel Castro from power. Eisenhower demanded that the CIA take extraordinary precautions to avoid direct evidence of U.S. involvement (Poznansky, 2020). But as CIA agents were secretly meeting Cuban contacts and building 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 (Eisenhower, 1960). The next month, the administration formally severed diplomatic ties with Cuba (DoS, 2023)—a wholly symbolic gesture, as the U.S. ambassador had already been recalled and communication between the governments already ceased entirely (ADST, 2023, p.53-59). Shortly after, New York Times reports began speculating that the CIA could be training and equipping an invasion force.¹

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 novel conceptualization of plausible deniability (Poznansky, 2022), focusing on the ability of audiences to draw inferences not solely on the basis of direct evidence, but relying on circumstantial evidence as well.

Consistent with existing research into World Order and covert action, we focus on a setting in which powerful states (we label them Interveners) are concerned with generating various forms of audience costs for taking overt military actions abroad—perhaps because those actions violate important norms or international laws which the Intervener claims to adhere to (Ikenberry, 2000; Bull, 2002). Interveners will thus 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).² We argue that international and domestic audiences do not only rely on direct evidence to determine whether or not a covert action has

¹NY Times, Jan 6 and Jan 14, 1961

²See Carnegie (2021) for a review.

taken place. Rather, they draw strategic inferences about whether a state authorized a covert 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.

To navigate the complex set of tradeoffs and informational asymmetries inherent to this strategic setting, we develop a game-theoretic model of foreign intervention and accountability. In the model, an Intervener has multiple policy levers available: one which is taken publicly, and commonly understood to be consistent with international norms; and another which is taken covertly, and in violation of those norms. The leader seeks to use these policy levers to achieve her desired outcomes, while maintaining a reputation among some domestic and/or international audience that she is the type of leader who values international norms.

If audiences draw strategic inferences, then plausible deniability is vastly more complicated than past secrecy scholarship has recognized. This consideration affects our understanding of 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)³—to penetrate US armored vehicles in Iraq?

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 actions 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

³Senate (2012)

to the accompanying public action, and less likely to suspect that the Intervener engaged in morally repugnant covert actions.

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 conditions under which Interveners will optimally pursue one approach over 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.

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 Interveners must employ the tactic sparingly so as to maintain unpredictability. Most surprisingly, we find that the frequency with which Interveners will deploy cover stories can be non-monotonic in the degree of heightened media scrutiny that their public actions invite: if covert action is not exposed despite the increased scrutiny, the audience is even more confident that no covert action took place—making cover stories more attractive the more scrutiny they bring. 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 Jacobo Arbenz 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 Administration sought to avoid direct evidence of US involvement through tight operational controls, and by distancing themselves publicly from the coup plotters as the coup was ongoing. Our analysis illuminates important events that are overlooked by existing accounts. We observe that the Administration and CIA planners believed that many in Latin America would blame the US for Arbenz’s removal even in the absence of direct evidence, because of strategic inferences. We provide evidence that highly publicized actions, including shipping embargoes, and sanctions and protests registered through the Organization of American States (OAS), were partly designed to disclaim

responsibility for the coup plot. We further show that after the mission was complete, the US government highlighted their public actions at the OAS as a cover story to disclaim responsibility for covert action.

Our theory holds direct implications for analyzing situations of conflict where attribution is ambiguous. Outside of international security, it also 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, N.d.; Terman and Byun, 2022); but critics still worry that any constraining effects are undermined by the ability of powerful states to exploit covert action (Carson, 2018). Our theory suggests a practical limitation on how frequently states can use covert action to circumvent international responsibilities, especially if they use covert actions repeatedly over time. Outside of international relations, our model contributes to the body of theoretical research on electoral accountability and adverse selection (Ashworth, 2012). This literature has rationalized counterintuitive behaviors such as showing-off (Gleason, 2017), admitting ignorance (Backus and Little, 2020), and adopting extreme ideological stances (Izzo, 2022). In a similar vein, we rationalize why leaders implement, then broadly publicize, ineffective and costly policies.

1 Secrecy, international reputation, and plausible deniability

In this section, we advance a conceptual innovation for research into secrecy. First, we review standard arguments that states fear backlash when pursuing coercive actions that are in violation of prevailing international laws and norms. Second, we argue that existing research narrowly construes plausible deniability. We argue that much insight can be gained into research on World Order, secrecy, 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 may worry that the Target will retaliate (Carson, 2015), or be concerned with drawing disapproval from third-party

audiences when coercive actions violate international norms and laws (Owen, 1994; Terman and Byun, 2022). As a result, when states take coercive actions that violate international laws and norms, they often do so in secret so as to avoid reputational costs and reprisals (Poznansky, 2019, 2020; Bull, 2002; Morse and Pratt, 2022).

While a large literature shares the view that powerful states are concerned with third-party audiences' reactions of their coercive actions abroad, the particular audience and the particular reactions of concern are context-dependent. It depends, in part, on the norms and laws to which the Intervener subscribes (Lake, Martin and Risse, 2021). For example, democracies that want to coerce fellow democracies, or meddle in another democracy's domestic affairs, turn to secrecy to avoid damaging their international reputation (Downes and Lilley, 2010; Reiter and Stam, 2002). States that want to commit human rights atrocities, or support an ally committing atrocities, often do so in secret because they are worried about backlash from the international community (Kremparic, 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 Intervener's specific national values (Clare, 2007); whereas regional actors, such as the members of the African Union, may be more concerned with other member governments violating sovereignty norms, no matter what their specific national values are. Also at issue is the extent to which laws and norms conflict (Farrell and Newman, 2021), and therefore the extent to which a state can craft a justification for its actions (Stein, 2000). For example, the decision to intervene against state-sponsored genocide pits a US commitment to uphold sovereignty against human rights (Finnemore, 2003).

The important point for our strategic theory is that in some contexts, certain coercive actions will invite severe backlash, but others will not. For brevity, we refer to coercive actions that an Intervener worries will cause backlash as **unscrupulous** actions. To illustrate, consider that concerns about the United States' reputation in Latin America, and the response of regional partners, partly motivated Congress to prohibit the Reagan administration from arming the Contras to oust Nicaraguan president Daniel Ortega. But Congress was less opposed to the use of economic sanctions or diplomatic isolation as means to influence Nicaraguan politics. In this case, the backlash that the administration was concerned with inciting was not associated with the policy *outcome* per se (ousting Ortega), but rather with the *means* by which the outcome was pursued. Our anal-

ysis takes as a starting point the fact that certain policy actions are commonly understood to be unscrupulous, while other actions are generally deemed more acceptable—even when targeting the same outcome as the unscrupulous actions—and we analyze the leader’s choice of policy actions given these perceptions.

1.2 Plausible deniability: A strategic theory

Consistent with existing work, our theory takes as given that audiences draw negative inferences about an Intervener upon observing direct evidence of covert intervention; it naturally follows that Interveners will take pains to avoid the revelation of such direct evidence of their unscrupulous behavior. Where we depart from previous research is in our assumptions regarding what it takes to sustain plausible deniability. In existing theories, whether plausible deniability is maintained is treated as a deterministic function of the direct evidence that is revealed. For instance, in the game-theoretic analysis most similar to ours, [Spaniel and Poznansky \(2018\)](#) assume that a cost is automatically imposed on the administration when covert action is revealed, but do not allow for the possibility of reputational costs arising from inference or speculation on the part of the audience. Such a focus on direct evidence is far-reaching throughout the literature. In a comprehensive review, [Poznansky \(2022, 523-524\)](#) identifies 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, and 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 much more likely 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 as to avoid culpability following the deaths of government 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 the outcomes that result from 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.

Yet there is only so much we can claim from an intuitive account. It is not immediately clear how leaders navigate the tradeoffs between avoiding direct evidence during mission planning, and building a long-term cover story. The problem is further complicated by the strategic feedback between the leader and audience: the leader knows the audience is drawing strategic inferences, and adjusts her behavior accordingly; meanwhile, the audience knows the leader is acting in anticipation of their response, and so adjusts their punishment/reward strategy accordingly. To gain a better understanding of this strategically complex interaction, we turn to our formal analysis.

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 in turn holds the agent accountable for his performance (see Miller, 2005). Within this context, the strategic challenge that interests us is adverse selection (Ashworth, 2012): the agent’s value to the principal depends on their level of quality along a privately-known attribute, and the principal’s goal is to discern the agent’s quality, retain the high-quality agents, and replace the low-quality agents. Researchers have examined a range of different agent attributes, including competence, honesty, and 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, 2024).

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 policy-making 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 the leader likewise believes to be against the public interest, but doing so in the hope that he is proven wrong (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-build our principal-agent model to meet the novel reputational concern described above. That is, we focus on the *scrupulousness* of the leader of the intervening state. This quality captures the extent to which a leader has internalized the international norms and institutional commitments that her state commits to uphold, or her willingness to violate them when it is politically expedient to do so.⁴ Holding fixed the nature of the foreign policy issue, the feasibility of different policy options, 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. 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. This particular reputational concern, along with the introduction of multiple (covert and public) policy levers, are significant formal departures from existing models of policy-making and pandering, and generate novel insights regarding leaders’ ability to leverage their private information and multiplicity of available policy tools in order to achieve their policy objectives and evade accountability.

⁴As 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 Table 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 $r \in \{0, 1\}$.

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 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 the Intervener's own electorate, or 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, and 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, $y \in \{0, 1\}$, according to the probabilities:

	$a_c\omega_c = 1$	$a_c\omega_c = 0$
$a_p\omega_p = 1$	α_{pc}	α_p
$a_p\omega_p = 0$	α_c	α_0

Table 1: Probability of policy success, $Pr(y = 1|a, \omega)$

Here, α_p (α_c) denotes the probability of success when only feasible public (covert) action is taken; α_{pc} is the probability of success when both feasible actions are taken; and α_0 denotes the baseline probability of success absent any intervention from L , or due to “random luck”. We assume $0 < \alpha_0 \leq \alpha_c \leq \alpha_p \leq \alpha_{pc} < 1$, meaning that either action, when feasible, increases the odds of policy success but does 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.⁵

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: we assume that L observes both ω_c and ω_p , while A holds prior beliefs $Pr(\omega_j = 1) = \tau_j$. Substantively, this means that the leader has a better (though still imperfect) understanding, as compared to the audience, of the likelihood that a given policy action will result in a successful outcome. Returning to the example of Eisenhower’s Castro policy, now-declassified documents indicate that the administration privately knew that the Soviets were willing and able to mitigate the impacts of both US oil embargoes⁶ and sugar quotas; as such, NIE 85–2–60 argued that “Fidel Castro will almost certainly remain in power through 1960”, despite the overt policies being pursued. Important for our argument is the notion that these intelligence assessments were *private*, and that some relevant foreign and domestic audiences faced uncertainty as to whether these sorts of policies could in fact contribute to the downfall of the Castro regime (CIA, 1960).

The leader’s second informational advantage over the audience pertains to the leader’s intrinsic

⁵Note 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 .

⁶See: Memorandum of Discussion at the 450th Meeting of the National Security Council, Washington, July 7, 1960

willingness to abide by international norms and institutions. We refer to this quality as the leader’s *scrupulousness*, 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 $\bar{\mu} \in (0, \pi)$: that is,

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

Intuitively, we can think of $\bar{\mu}$ as the “reservation payoff” that A receives from foregoing future cooperation with L (as an international audience), or from withdrawing political support from L (as a domestic audience); more directly, it determines the level of suspicion that A would have to hold regarding L ’s unscrupulousness before being willing to impose punishment.

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 $\delta \geq 0$ 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 scrupulousness $\mu^{a_p, y, z} = Pr(\theta = 1 | a_p, y, z)$. In equilibrium, A punishes if $\mu^{a_p, y, z}$ falls below the 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,

$$U_L(a) = y - a_p k_p - a_c k_c^\theta + r\beta \quad (2)$$

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,⁷ and for shorthand we will denote $k_c = k_c^0$. We analyze Perfect Bayesian Equilibria of the model.⁸

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. Specifically, in the equilibrium of interest, the leader pursues public instead of covert action when public action is feasible; and when both covert and public actions are infeasible, the leader refrains from taking either action. (This strategy is depicted in Table 2.) It is under the condition that covert action is feasible, but public action is not—that is, when $\omega_p = 0$ and $\omega_c = 1$ (which are privately known to the leader)—that the unscrupulous leader faces a dilemma. In this situation, he must choose between advancing a foreign policy objective through unscrupulous means, or forgoing his preferred policy outcome in order to avoid reputational harm and resulting 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 (or employs a “cover story”) when he also takes an infeasible public action ($a_p = 1$ when $\omega_p = 0$).*

⁷It suffices to assume that $k_c^1 > 1 + \beta$.

⁸The analysis proceeds under a number of technical assumptions stated in the appendix, in the form of parameter restrictions and equilibrium selection rules which serve to focus our attention on the equilibrium characterized below.

		Covert action feasibility	
		$\omega_c = 1$	$\omega_c = 0$
Public action feasibility	$\omega_p = 1$	Public action only	
	$\omega_p = 0$	$\sigma^1(\omega) = (0, 0)$ $\sigma^0(\omega) = (\hat{\sigma}_c, a_c \hat{\sigma}_p)$	No action

Table 2: Leader’s equilibrium strategy, in the equilibrium selected by Assumption 3 (see appendix)

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 *Under Assumption 1 (parameter restrictions) and Assumptions 2 and 3 (equilibrium refinements), 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 unscrupulous 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.*

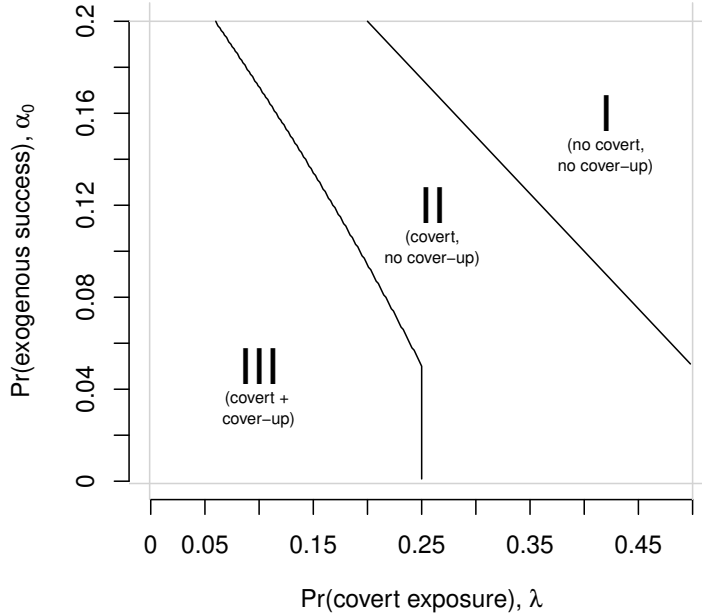
There exist thresholds λ', λ'' , where $\lambda' \leq \lambda'' = \frac{\alpha_c - \alpha_0 - k_c}{\beta}$, such that:

- *When $\lambda > \lambda''$, the equilibrium is Class I.*
- *When $\lambda \in (\lambda', \lambda'')$, the equilibrium is Class II.*
- *When $\lambda < \lambda'$, the equilibrium is Class III.*

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, so 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.⁹ The only risk the leader faces in this case is that of direct exposure; if the

⁹Note that both Class II and Class III equilibria include conditions in which the leader plays either a mixed or pure covert action strategy, 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 direct exposure is too high, so the leader avoids covert action entirely.
II	Yes	No	$\lambda \in (\lambda', \lambda'')$	The risk of direct exposure is low enough that the leader takes covert action; but high enough that the absence of direct evidence reassures the audience that no covert action occurred.
III	Yes	Yes	$\lambda < \lambda'$	The risk of direct exposure is so low that the absence of direct evidence of covert action is uninformative. The audience will blame the leader for policy success on the basis of circumstantial evidence. So the leader uses performative public action as a cover story.

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 serves 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.

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.

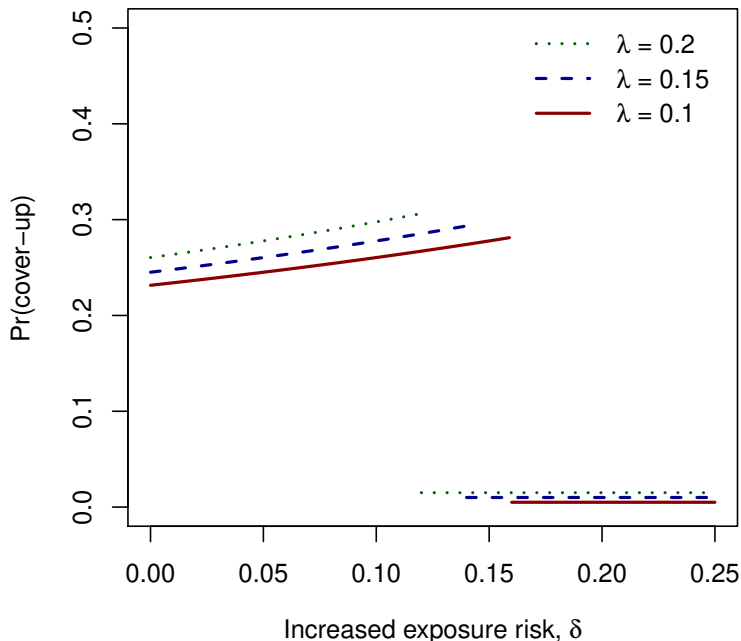
There are, however, important limitations on the leader’s ability to skirt accountability through cover-up actions. First, and most fundamentally, the use of cover stories is partially self-defeating: when the audience expects the leader to use cover stories, they become increasingly skeptical of successful 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 story, 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 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. When we increase δ from zero (starting on the lefthand side of the figure and moving rightward), there are direct and indirect effects that push in competing directions. First, as a direct effect, there is an increased risk that the leader’s public action draws attention which causes his covert action to be exposed; this serves to disincentivize the use of cover-up action. But second, as an indirect effect, the audience’s inference upon observing public action *without* covert revelation becomes more favorable; this serves to incentivize the use cover-up action. The left half of Figure 3 depicts an equilibrium in which the latter effect dominates.¹⁰ 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

¹⁰Note that this equilibrium is not unique; under these same conditions, there also exists an equilibrium in which $\Pr(\text{cover-up})$ is constant in δ , up until the discontinuity depicted in the figure. See Corollary 2 in the appendix for a formal statement of this comparative static result.

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 $\Pr(\text{cover-up})$ 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: $\tau_p = 0.5, \tau_c = 0.8, \alpha_p = 0.8, \alpha_c = 0.4, \alpha_0 = 0.05, \kappa_p = 0.1, \kappa_c = 0.1, \beta = 0.5, \pi = 0.5, \bar{\mu} = 0.4$.

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. Beyond these established results, 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 formed 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 ($\alpha_p > \alpha_0$). 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 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 marked the first time in Guatemala’s history that power was peacefully transferred from one democratically-elected leader to another. From an institutional perspective, the 1950 election suggested that democracy was working in Guatemala (Fraser, 2005, 487). But it was not working for the United States. Answering the calls of the Guatemalan communist party, newly-elected President Jacobo Arbenz implemented extensive land and agrarian reforms (Schlesinger and Kinzer, 1982, 53), which directly challenged US commercial and political interests. US policymakers were also concerned by the number of communists appointed to government positions (Immerman, 1982, 108). In his memoirs, Eisenhower worried that a major threat to his objectives was that “Communism was striving to establish its first beachhead in the Americas by gaining control of Guatemala.”¹¹

In August, 1953, Eisenhower authorized the covert CIA operation PBSUCCESS. The first phase of the operation involved establishing bases in neighboring countries, which would be used

¹¹Quoted in Schmitz (1999, 179)

to train and arm 480 Guatemalans to overthrow the Arbenz government. The CIA also groomed a staunch 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 involved offensive psychological operations aimed at convincing 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 psychological operations had already begun to undermine widespread confidence in the Arbenz government.

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

Following best practices in the evaluation of formal models, we use this case to illustrate the empirical plausibility of our theory. In section 3.1 we detail our case selection methodology (Bates, 1998). Then, following Goemans and Spaniel (2016), Joseph, Poznansky and Spaniel (2022) and others, we evaluate our theory by examining primary evidence of Eisenhower’s decision-making processes, paying particular attention to the choice nodes that we model. We develop case-specific hypotheses about what our theory expects to observe, and evaluate them against the leading alternatives. Section 3.2 asks whether the Eisenhower Administration worried about strategic inferences as they implemented their covert actions. Section 3.3 details the overt actions the Administration took, and explores how well existing accounts can explain these actions. Section 3.4 provides evidence that the cover story mechanism can explain these overt actions. Finally, section 3.5 addresses concerns and alternative explanations.

3.1 Case selection, and calibrating the parameters

Following Bates (1998), our main concern was finding a case in which the leader of the intervening state faced the core strategic tension characterized by our model, and the initial conditions fit the parameters that support the cover story equilibrium. Below, we first discuss why the Eisenhower administration’s overarching objective of toppling communist-leaning but popular governments fits

the broad contours of our model. We then discuss why the Guatemala case in particular closely matches the conditions for the cover story equilibrium.

3.1.1 Selection of the Eisenhower administration

At the broadest level, our theory applies to a major world power that is both interested in shaping political developments abroad, and concerned with maintaining a reputation among various audiences for only doing so through scrupulous means. This makes the United States in the latter half of the 20th century a natural choice. The early Cold War, and the Eisenhower administration in particular, are especially well fitting on both of these counts.

Eisenhower’s primary foreign policy objective was to stop the spread of communism (Schmitz, 1999). To win over the developing world, the US’s overarching strategy was to promote the principles of sovereignty, self-determination and democracy as core tenants of the Liberal Order (Rabe, 1988, 166). Yet there remained uncertainty across the developing world as to the US’s true commitment to the ideals it espoused—reflecting the core uncertainty in our model regarding the leader’s true “type”. Eisenhower understood that using military power to overturn a democratically elected government would reveal him as highly unscrupulous, in the sense implied by our theory (Poznan-sky, 2019, 86). In the Guatemala case specifically, the Administration estimated that such overt disregard for liberal principles would lead to severe “international consequences.”¹² Thus the core tension of wanting to shape political developments abroad, while avoiding the reputational damage that would follow from doing so through unscrupulous means, is a prominent tension faced by the Eisenhower administration in this early Cold War period.

There are two potential concerns worth addressing regarding the selection of the Eisenhower administration as the general setting of our case study. First, some historical accounts question the extent of President Eisenhower’s direct involvement in foreign policy choices, and the degree to which administration policies actually reflected his own worldview (Divine, 1981); this work instead suggests that key advisers, notably the Dulles Brothers (with Allen Dulles as CIA director, and John Foster Dulles as Secretary of State), played an outsize role. Yet other work argues Eisenhower was more skillful and directly involved in policy decisions (see McAuliffe, 1981). Recognizing that this debate exists, we analyze documents that provide insights into the reasoning of the Administration

¹²See Memorandum for Col J. C. King, PBSUCCESS 20th Jan 1954.

as a whole—including Eisenhower and the Dulles brothers, as well as their appointed delegates within the CIA’s Directorate of Plans and the Guatemalan Embassy.

Second, some might worry that the CIA was unusually popular with the US public in 1954, and therefore the US public would have ignored even direct evidence of a covert operation. However, support was unusually strong only because “American people remained in blissful ignorance of the CIA’s covert objectives” (Jeffreys-Jones, 2022). If those actions were exposed, opinions may have changed. Further, in the case we analyze, administration officials primarily expressed concern over international audiences (including audiences within Guatemala, as well as throughout Latin America and beyond) rather than U.S. domestic audiences when discussing the risk of exposure for PBSUCCESS (Schmitz, 1999). The relative concern over different audiences may shift across different cases and different time periods, and we believe our theory can accommodate this variation.

3.1.2 Selection of the Guatemala intervention

Guatemala is an especially important case to examine (Bates (1998)’s second criterion) because it represented the first major communist foothold in the Americas. For our purposes, this also means that the value of policy success (normalized to 1 in our model) is high, relative to the direct costs of action (k_p and k_c) and the reputational concerns (β). The analytical clarity of this particular case, relative to other regime change operations that Eisenhower authorized and pursued, is also aided by the fact that all the salient choices were made within the Eisenhower administration.¹³

Consistent with our model, in pursuing regime change in Guatemala, the Eisenhower administration faced policy options that can largely be characterized as either public and scrupulous, or covert and unscrupulous. Concerned that brazen military intervention into a regional democracy would sour opinions of the US throughout Latin America (Schmitz, 1999, p181), Eisenhower only seriously considered military actions that could be undertaken covertly. By contrast, economic sanctions, or public diplomacy that was designed to expose the failures of communism and cause domestic unrest, were not seen as inconsistent with Liberal Order, and thus more tolerable to foreign and domestic audiences.¹⁴

¹³As another potential case, many features of Eisenhower’s reasoning to oust Castro fit our cover story equilibrium. But the case is more complicated to analyze because the decisions spanned multiple administrations, with Kennedy ultimately approving the mission. For an interesting overview of cover story references in this case see CIA, Official History of the Bay of Pigs Operation, V II, pp12-14.

¹⁴See Memorandum for Col J. C. King, PBSUCCESS 20th Jan 1954.

The main condition that our model identifies for the cover story equilibrium to hold is that the transparency parameter, λ , is low. At this time, CIA was optimistic that they could successfully orchestrate a coup and conceal direct evidence of their involvement. Their confidence was heightened after using the same playbook to oust Mossadeq in Iran undetected (Cullather, 2006, 7). The state of information technology in Guatemala in 1954 also meant the chance of direct exposure was low (Joseph and Poznansky, 2018). Further, in this period, it was not widely known whether the CIA was even in the business of orchestrating regime change, or if so, with what methods. Thus, even if some audiences sought out direct evidence of US covert involvement, they may not have even known what to look for.

The cover story equilibrium also requires that the likelihood of a favorable outcome absent intervention (that is, the α_0 parameter) is low. In this case, Eisenhower firmly believed that if he did not act, that communism would expand across the developing world (Schmitz, 1999, 179). Eisenhower did believe 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 because of Soviet meddling, were not left to their own devices. As a result, he was gravely concerned that absent US intervention, communist parties would take hold in much of the developing world (Schmitz, 1999, 182).

3.2 Plausible deniability was difficult but important

From the outset of planning Operation PBSUCCESS, plausible deniability was viewed as essential to the mission's success. A recurring reminder from administration officials to mission planners was: "don't get caught" (FRUS, 1954). Consistent with existing theoretical arguments (Joseph and Poznansky, 2018), this included diligent efforts to avoid direct evidence of US involvement. According to Immerman (1982, 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).

3.3 The Puzzle of Overt Action

Given the intense focus on maintaining secrecy, we might expect that the administration 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 next day (Cullather, 2006, 80). (Ironically, the CIA had planned to fabricate a Soviet arms cache, under operation WASHTUB (Cullather, 2006, 101), which the US would then discover and exploit publicly; the fortuitous discovery of an actual weapons shipment obviated the need for this particular ploy.) Rather than minimize the episode, Eisenhower expressed public outrage: he invoked the Monroe Doctrine, which called for the exclusive influence of the United States in Latin America, and proceeded to impose a naval blockade to prevent future arms shipments into Guatemala (Cullather, 2006, 79).¹⁵

At the same time, the US convened an emergency meeting of the Organization of American States in which Dulles delivered an impassioned speech attacking the Guatemala government. This was at Eisenhower’s direction, who instructed his diplomats that “By every proper and effective

¹⁵By our theory, the blockade constituted a “public” action, which was not deemed to be “unscrupulous”. In this instance, some Latin American states viewed the blockade as an unjustified violation of sovereignty; but the backlash was relatively minor (Friedman, 2010, 672), particularly compared to the backlash that would have followed from the revelation of the U.S.’s even more unscrupulous covert activities. Our model does allow for (and indeed, requires) public actions to carry some cost, k_p , which is non-zero but much smaller than the cost that scrupulous types incur from covert action, $k_c^{\theta=1}$.

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”, Eisenhower 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).

This argument is not inconsistent with our theory.¹⁶ We account for the possibility that overt and covert policies can raise the chance of success simultaneously. The important question for our analysis is: does Eisenhower’s desire to deploy all policy tools fully account for the extent and kind of overt action that we observe? If the answer is yes, we would expect that the administration to only publicize overt policies when doing so confers some operational advantage. We believe that two aspects of how Eisenhower publicized overt actions do not fit this explanation.

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 PBSUCCESS was commanding operations from an undisclosed location in Florida (codenamed LINCOLN). The more attention within the United States, the more media scrutiny would follow, raising the chance of exposure at this critical operational moment. Further, 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 [overt] action... [that were] altogether infeasible” (CIA, 1953).

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

¹⁶Cullather’s historical analysis implies that foreign audiences perceived US overt actions as plausibly effective. Indeed, international messages can facilitate local actors in coordinating regime change (Little, 2017; Malis and Smith, 2019, 2021).

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” (CIA, 1954*h*). So 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.

Altogether, these fact patterns point to performative public actions taken by Eisenhower administration officials which provided no discernible operational advantages—and may have even undermined operational security—and thus cannot be explained by a standard logic of using all available and effective means to achieve the desired policy objective.

3.4 The cover story explanation

As outlined above, our theoretical model carries three key observable implications: first, that U.S. decision-makers will exhibit concern for strategic inferences made by audiences, even in the absence of direct evidence of wrongdoing; second, that they will pursue public actions which are performative but privately understood to be ineffective; and third, they will seek to credit those public actions with the policy successes which were actually brought about by the unscrupulous covert action. We discuss each of these in turn.

While planning PBSUCCESS, administration officials expressed an acute concern for strategic inferences. 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” (FRUS, 1953) 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.

CIA Deputy Director for Plans Frank Wisner laid out the concern even more explicitly. Wisner cautioned 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” (CIA, 1954*f*). He went on to warn: “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. But it is equally

true that they would see the hand of the U.S. in any uprising whether or not sponsored by the U.S.” (CIA, 1954e). By the logic of our theory, Wisner is articulating the view that, given a low level of transparency (low λ), the *absence of evidence* of U.S. involvement does not provide sufficiently compelling *evidence of absence* of U.S. involvement in order to avoid blame for the observed outcome.

Given these concerns, Wisner and his staff took an active role in crafting cover stories to allay suspicions among the target audience. In a discussion about how to prevent Latin American audiences from speculating of US involvement, Wisner argued that “it might be a good idea to cry wolf several times before D-Day” (CIA, 1954g). In June, Wisner’s subordinates managing operations from LINCOLN observed with disappointment that US Ambassadors in Honduras and Guatemala were not publicly voicing the US position. Even though they did not think diplomatic statements would affect whether or not the coup prevailed, they still thought it was “essential that for diplomatic battle the hole created by non-participation should be filled” (CIA, 1954d).

We further expect that upon employing these cover stories, administration officials should seek to explicitly connect the observed outcomes to the public actions, so as to disclaim any covert regime change operations. This is a hard test, because political elites rarely document the reasoning behind what they say. However, in two instances that we identified, elites explicitly documented a cover-story motivation for their discussion of US overt policies. For example, Second Secretary Hill of the US Embassy in Guatemala recounted his conversation with a Guatemalan political elite (whose name is still classified)¹⁷ as follows:

I told [redacted] that Ambassador Patterson had been quite correct in pointing out the US policy of non-intervention... but [redacted] 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. ...*In talking in this vein to [redacted] it was my intention to give him the impression that the US had no concrete plan for intervention in*

¹⁷The context suggests the unknown subject was influential in Guatemalan politics, not currently in government, and somewhat concerned about the communist trajectory, but not loyal to the United States.

the domestic affairs of Guatemala and continued its non-intervention policy. (emphasis added)¹⁸

This last sentence directly describes the logic of our argument: the reason that Hill highlights overt policies is to disclaim covert policies.

Finally, after the completion of PBSUCCESS, we find that US officials continued to publicize the overt actions taken in order to cover up the covert actions. 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 Latin 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 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 (CIA, 1955).

As the quote shows, the story being put forward by the administration is that the US pursued its objectives in Guatemala through the OAS; by implication, the US would not have been involved in other, unscrupulous means of pursuing the same objectives.

What is more, it appears that analysts who studied US actions at the time refer to these actions to offset suspicions of covert involvement. For example, 2 years after Arbenz was ousted, Taylor

¹⁸CIA (1954b) For a second example with a different unknown subject, see CIA (1954a).

(1956) published a comprehensive “Critique of United States Foreign Policy” surrounding Arbenz’s removal. He comprehensively reviewed journalistic inquiries into US policies, and academic and policy investigations into the US role published across Latin America and the United States. Taylor shows that a handful of pundits 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 focused 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.

Altogether, this evidence shows that the executive was concerned about strategic inferences; and 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 the administration referred back to these overt policies to divert attention away from their sponsorship in the years after the coup succeeded.

3.5 Clarifications

Following Bates (1998) and others, we narrowly focused on the aspects of the case that are most salient for the decision nodes in our theory. We now address three potential concerns from researchers who study PBSUCCESS from a broader range of historical perspectives.

First, our goal was to show that the cover story mechanism was one overlooked way that mission planners avoided strategic inferences. We accept that it was not the only way. 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.¹⁹ The CIA also

¹⁹There 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.”

armed the coup-plotters with weapons it purchased from the Dominican Republic to implicate them (CIA, 1954c). We do not think plausible deniability is uni-causal, and 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, “Added support in cloaking the U.S. hand exists in the number of other countries”—such as Nicaragua and Honduras, where the foreign training bases were located—“which both have good reasons for wanting to see the replacement of the Arbenz Government and have the means for backing a coup of the size planned.”²⁰

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 offset suspicion just enough so that the Intervener can avoid backlash. Consistent with our theory, some suspicion did arise. But that suspicion was not widespread (see discussion of Taylor, 1956, above). Indeed, historians who consider sustaining plausible deniability as part of mission success, argue that PBSUCCESS was a successful case of covert regime change (Immerman, 1982; Schmitz, 1999).

Third, one might wonder if 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 alongside 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 when international and domestic audiences form perceptions and understandings of international affairs, they draw inferences about the actions of powerful states not only from direct evidence, but also from circumstantial evidence—the context, interests, capabilities, and

²⁰<https://history.state.gov/historicaldocuments/frus1952-54Guat/d133>

policy outcomes. 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 occur in the absence of intervention, a state can avoid attribution of their covert action through the use of a cover story. The ideal cover story is more consistent with international laws and norms than the secret policy that the leader wants to disclaim responsibility for, and constitutes a plausible explanation for observed policy outcomes. Cover stories are rational mechanisms to deny morally repugnant secret policies, even if they invite additional scrutiny that raises the risk that such policies are exposed directly, and even if they are somewhat unpopular among international or domestic 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 remove Arbenz as president of Guatemala, supports our theory. We find evidence that the Eisenhower Administration worried about strategic inferences, believing that many in the region would assume the US was responsible even if no direct evidence emerged of CIA involvement. They deployed a range of tactics to avoid attribution for Arbenz's ultimate ouster. To this end, one important but overlooked mechanism was a series of performative overt policies. We find direct evidence that Embassy staff in Guatemala discussed US public actions to disclaim the possibility of covert regime change operations. We also show that in the years after Arbenz resigns, the administration points to these open policies to account for its own involvement in Arbenz's removal, and to help discredit speculation that a nefarious secret policy was at work.

Our theory holds two policy implications. First, it helps public accountability activists appreciate that building extensive monitoring capabilities may, in some cases, work against their objectives. 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 from an absence of evidence that no unscrupulous policy took place. This, in turn, may make covert action more attractive.

Second, the logic of strategic inferences means that grey zone conflict may play a different role in great power competition—particularly, the emerging US-China competition—than existing

research expects. Existing work emphasizes that states can avoid escalation and retaliation if they can conceal direct evidence of an attack. But the United States often employs secrecy to maintain its reputation as compliant with liberal values, while pursuing goals that clearly violate those principles. To the extent that strategic inferences are plausible, grey zone attacks may not be viable; and this constraint may be more binding on the US, as the power most expressly concerned with maintaining the principles of the Liberal Order. This may present an autocratic advantage in the use of grey zone conflict to influence third parties in the decades to come. Our theory also shows that this disadvantage can be partially offset via a cover story.

Finally, our cover story mechanism likely holds theoretical implications beyond the particular setting of a state concealing its military operations for the sake of feigning compliance with its espoused values. Indeed, it could apply in any setting where a leader can independently deploy multiple policies secretly or overtly, and has a specific kind of reputation that they hope to uphold. For example, it is well known that Kennedy made a secret (and, if exposed, highly unpopular) promise to withdraw missiles from Turkey in exchange for Soviet compliance during the Cuban missile crisis, and that Kennedy's decision to publicize other policies brought additional scrutiny from analysts. These publicized overt policies could be interpreted as a cover story, to deflect suspicions that the desired outcome was achieved through less palatable means which were out of public view.

References

- ADST. 2023. "Oral Histories Project, Cuba." Association for Diplomatic Studies and Training.
- Ashworth, Scott. 2012. "Electoral Accountability: Recent Theoretical and Empirical Work." *Annual Review of Political Science* 15:183–201.
- Ashworth, Scott and Ethan 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. "Timing of cyber conflict." pp. 1298–1303.

- Backus, Matthew and Andrew T. Little. 2020. "I Don't Know." *American Political Science Review* 114(3):724–743.
- Baliga, Sandeep, Ethan Bueno de Mesquita and Alexandor Wolitzky. 2020. "Deterrence with Imperfect Attribution." *American Political Science Review* 114:1155–1178.
- Bates, Robert H. 1998. *Analytic narratives*. Princeton University Press.
- Bowen, Gordon L. 1983. "U.S. Foreign Policy toward Radical Change: Covert Operations in Guatemala, 1950-1954." *Latin American Perspectives* 10(1):88–102.
- Bull, Hedley. 2002. *The anarchical society : a study of order in world politics*. Columbia University Press.
- Canes-Wrone, Brandice, Michael C. Herron and Kenneth W. Shotts. 2001. "Leadership and Pandering: A Theory of Executive Policymaking." *American Journal of Political Science* 45:532.
- Carnegie, Allison. 2021. "Secrecy in International Relations and Foreign Policy." *Annual Review of Political Science* 24:213–233.
- Carnegie, Allison and Austin Carson. 2018. "The Spotlight's Harsh Glare: Rethinking Publicity and International Order." *International Organization* 72(03):627–657.
- Carson, Austin. 2015. "Facing Off and Saving Face : Covert Intervention and Escalation Management in the Korean War." *International Organization* Forthcomin.
- Carson, Austin. 2018. *Secret wars : covert conflict in international politics*.
- CIA. 1953. "Wisner's Memorandum for Chief, Western Hemisphere Division (CREST, No. 0000914589).". https://www.cia.gov/readingroom/docs/DOC_0000914589.pdf.
- CIA. 1954a. "Conversation Between U.S. Ambassador Hill and SA/DD/P, 25 Jan (CREST no. 0000914014).". https://www.cia.gov/readingroom/docs/DOC_0000914014.pdf.
- CIA. 1954b. "Dispatch No. HGG-A-619. Memorandum of Conversation, From COS, Guatemala to Chief WHD (CREST DOC'0000914259).". https://www.cia.gov/readingroom/docs/DOC_0000914259.pdf.

- CIA. 1954c. “Memorandum for Goodbourne, Progress Report - PBSUCCESS. 11-17 May (CREST No. 0000923551).”.
- CIA. 1954d. “Memorandum from LINCOLN to RYBAT (CIA) Director, 21 June, 1954.” *Central Intelligence Agency, Job 79-01025A, Box 9, Folder 2.* .
- CIA. 1954e. “Memorandum From the Deputy Director for Plans of the Central Intelligence Agency (Wisner) to Director of Central Intelligence Dulles.” *Central Intelligence Agency, Job 79-01025A, Box 151, Folder 6. Top Secret.* .
- CIA. 1954f. “Minutes of Weekly PBSUCCESS Meeting. 20 April, 1954 (CREST No. 0000916252).” .
https://www.cia.gov/readingroom/docs/DOC_0000916252.pdf.
- CIA. 1954g. “PBSUCCESS Weekly Meeting, 30 March 1954 (Crest No. 0000916849).” . https://www.cia.gov/readingroom/docs/DOC_0000916849.pdf.
- CIA. 1954h. “Synthesis of [classified]’s remarks relevant to PBSUCCESS. April 22, 1954 (CREST no. DOC`0000135899).” . https://www.cia.gov/readingroom/docs/DOC_0000135899.pdf.
- CIA. 1955. “Progress Report on NSC 5432/1 (Crest No. CIA-RDP80R01731R003000030008-2).”.
- CIA. 1960. “NIE 85-2-60, the Situation in Cuba (CREST no. 0000132646).” . https://www.cia.gov/readingroom/docs/DOC_0000132646.pdf.
- Clare, Joe. 2007. “Domestic Audiences and Strategic Interests.” *The Journal of Politics* 69:732–745.
- Coe, Andrew J. 2018. “Containing Rogues: A Theory of Asymmetric Arming.” pp. 1197–1210.
- Cullather, N. 2006. *Secret History: The CIA’s Classified Account of Its Operations in Guatemala, 1952-1954*. Stanford University Press.
URL: <https://books.google.com/books?id=eDsG-0-p15sC>
- Divine, R. A. 1981. *Eisenhower and the Cold War*. Oxford University Press.
- DoS. 2023. “A Guide to the United States’ History of Recognition, Diplomatic, and Consular Relations, by Country, since 1776: Cuba.” . <https://history.state.gov/countries/cuba>.

- Downes, Alexander B. and Mary Lauren Lilley. 2010. "Overt Peace, Covert War?: Covert Intervention and the Democratic Peace." *Security Studies* 19(2):266–306.
- Eisenhower, Dwight. 1960. "Statement by the President Upon Issuing Proclamation Fixing the Cuban Sugar Quota at Zero." .
URL: <https://www.presidency.ucsb.edu/documents/statement-the-president-upon-issuing-proclamation-fixing-the-cuban-sugar-quota-zero>
- Fang, Songying. 2008. "The informational role of international institutions and domestic politics." *American Journal of Political Science* 52(2):304–321.
- Farrell, Henry and Abraham L. Newman. 2021. "The Janus Face of the Liberal International Information Order: When Global Institutions Are Self-Undermining." *International Organization* 75:333–358.
- Finnemore, Martha. 2003. *The Purpose of Intervention*. Cornell University Press chapter Changing Norms of Humanitarian Intervention, pp. 52–84.
- Fox, Justin and Kenneth W. Shotts. 2009. "Delegates or Trustees? A Theory of Political Accountability." *The Journal of Politics* 71(4):1225–1237.
- Fraser, Andrew. 2005. "Architecture of a broken dream: The CIA and Guatemala, 1952–54." *Intelligence and National Security* 20(3):486–508.
- Friedman, Max Paul. 2010. "Fracas in Caracas: Latin American Diplomatic Resistance to United States Intervention in Guatemala in 1954." *Diplomacy & Statecraft* 21(4):669–689.
- FRUS. 1953. "Draft Memorandum for the Record: Program For PBSUCCESS, November 12, 1953." . <https://history.state.gov/historicaldocuments/frus1952-54Guat/d65>.
- FRUS. 1954. "Memorandum for the Record. March 14." . <https://history.state.gov/historicaldocuments/frus1952-54Guat/d116>.
- Gleason, Judd. 2017. "Showing Off: Promise and Peril in Unilateral Policymaking." *Quarterly Journal of Political Science* 12(2):241–268.
- Goemans, Hein and William Spaniel. 2016. "Multimethod Research:." *Security Studies* 25(1):25–33.

- Grant, Donald. 1955. "Guatemala and the United States' Foreign Policy." *Journal of International Affairs* 9:64–72.
- Hakakian, Roya. 2011. *Assassins of the Turquoise Palace*. Grove Press.
- Houghton, Vince. 2019. *Nuking the Moon: And Other Intelligence Schemes and Military Plots Left on the Drawing Board*. Penguin Books.
- Huth, Paul K., Sarah E. Croco and Benjamin J. Appel. N.d. "Does International Law Promote the Peaceful Settlement of International Disputes? Evidence from the Study of Territorial Conflicts since 1945." . Forthcoming.
- Ikenberry, G. John. 2000. "After victory : institutions, strategic restraint, and the rebuilding of order after major wars." p. 293.
- Immerman, Richard. 1982. *The CIA in Guatemala*. University of Texas Press.
- Izzo, Federica. 2022. "Ideology for the Future." *American Political Science Review* pp. 1–16.
- Jeffreys-Jones, Rhodri. 2022. *Covert Action in the 1950s*. Oxford University Press Oxford pp. 33–48.
- Joseph, Michael F and Michael Poznansky. 2018. "Media technology, covert action, and the politics of exposure." *Journal of Peace Research* 55:320–335.
- Joseph, Michael F., Michael Poznansky and William Spaniel. 2022. "Shooting the Messenger: The Challenge of National Security Whistleblowing." *The Journal of Politics* 84(2):846–860.
- Krcmaric, Daniel. 2019. "Information, Secrecy, and Civilian Targeting." *International Studies Quarterly* 63:322–333.
- Lake, David A., Lisa L. Martin and Thomas Risse. 2021. "Challenges to the Liberal Order: Reflections on International Organization." *International Organization* 75:225–257.
- Levin, Dov H. 2021. *Meddling in the ballot box : the causes and effects of partisan electoral interventions*.
- Lindsay, Jon R. 2013. "Stuxnet and the Limits of Cyber Warfare." *Security Studies* 22:365–404.

- Little, Andrew T. 2017. "Coordination, Learning, and Coups." *Journal of Conflict Resolution* 61:204–234.
- Malis, Matt. 2024. "Foreign Policy Appointments.". Working paper. <http://mattmalis.github.io/research>.
- Malis, Matt and Alastair Smith. 2019. "A global game of diplomacy." *Journal of Theoretical Politics* 31(4):480–506.
- Malis, Matt and Alastair Smith. 2021. "State visits and leader survival." *American Journal of Political Science* 65(1):241–256.
- Maskin, Eric and Jean Tirole. 2001. "Markov perfect equilibrium: I. Observable actions." *Journal of Economic Theory* 100(2):191–219.
- Maskin, Eric and Jean Tirole. 2004. "The politician and the judge: Accountability in government." *American Economic Review* 94(4):1034–1054.
- McAuliffe, Mary S. 1981. "Eisenhower, the President." *The Journal of American History* 68(3):625.
- Miller, Gary J. 2005. "The Political Evolution of Principal-Agent Models." *Annual Review of Political Science* 8:203–225.
- Morse, Julia C. and Tyler Pratt. 2022. "Strategies of Contestation: International Law, Domestic Audiences, and Image Management." *The Journal of Politics* 84(4):2080–2093.
- Owen, John M. 1994. "How Liberalism Produces Democratic Peace." *International Security* 19:87.
- Poznansky, Michael. 2019. "Feigning Compliance: Covert Action and International Law." *International Studies Quarterly* .
- Poznansky, Michael. 2020. *In the shadow of international law : secrecy and regime change in the postwar world*.
- Poznansky, Michael. 2022. "Revisiting plausible deniability." *Journal of Strategic Studies* 45:511–533.

- Rabe, Stephen G. 1988. "Eisenhower and Latin America : the foreign policy of anticommunism." p. 237.
- Ramsay, Kristopher W. 2004. "Politics at the water's edge: Crisis bargaining and electoral competition." *Journal of Conflict Resolution* 48(4):459–486.
- Reiter, Dan and Allan 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, David F. 1999. *Thank God They're on Our Side: The United States and Right-wing Dictatorships, 1921-65*. University of North Carolina Press.
- Schram, Peter. 2022. "When Capabilities Backfire: How Improved Hassling Capabilities Produce Worse Outcomes."
- Schultz, Kenneth A. 2005. "The politics of risking peace: Do hawks or doves deliver the olive branch?" *International Organization* 59(1):1–38.
- Senate. 2012. "Iran's Support for Terrorism in the Middle East (S. Hrg. 112-606)." <https://www.govinfo.gov/content/pkg/CHRG-112shrg76693/html/CHRG-112shrg76693.htm>.
- Smith, Alastair. 1998. "International crises and domestic politics." *American Political Science Review* 92(3):623–638.
- Spaniel, William and Michael Poznansky. 2018. "Credible Commitment in Covert Affairs." pp. 668–681.
- Stein, Arthur. 2000. "The justifying State: Why Anarchy Doesn't Mean No Excuses."
- Taylor, Philip B. 1956. "The Guatemalan Affair: A Critique of United States Foreign Policy." *American Political Science Review* 50:787–806.
- Terman, Rochelle and Joshua Byun. 2022. "Punishment and Politicization in the International Human Rights Regime." *American Political Science Review* 116:385–402.

Appendix

Table of Contents

A Formal Appendix	1
A.1 Equilibrium Existence	2
A.2 Comparative Statics	8

A Formal Appendix

The goal of the formal analysis is to identify the conditions under which leaders pursue foreign policy objectives through unscrupulous covert means, and to explain when and how they use seemingly scrupulous public actions to provide a cover story and evade culpability for their unscrupulous covert behavior.

To focus attention on these strategic aspects of the situation, we will abstract away from others. In particular, we take as a starting point in the analysis that there exist some set of policy actions available to the leader which, if feasible, are largely unproblematic in legal, moral, and political terms. (We refer to these as “public” actions.) We wish to focus on the scenario in which the leader takes these actions when they are feasible, and pursues unscrupulous covert action only when the available public actions are not expected to succeed on their own. This restriction is accomplished by the equilibrium selection rule imposed in Assumption 3. The set of parameter restrictions imposed in Assumption 1 are sufficient to ensure that such an equilibrium exists. Assumption 2 serves to rule out other equilibria which are (potentially) technically viable but substantively uninteresting.

Notation:

- Let $\omega = (\omega_c, \omega_p)$; let $\sigma_j^\theta(\omega) = Pr(a_j = 1|\omega, \theta)$ for $j = c, p$; and let

$$\sigma^\theta(\omega) = (\sigma_c^\theta(\omega), \sigma_p^\theta(\omega)) = (\sigma_c^\theta(\omega_c, \omega_p), \sigma_p^\theta(\omega_c, \omega_p))$$

- 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, $\sigma_p^\theta(\omega; a_c) = Pr(a_p = 1|\omega, a_c, \theta)$.

Assumption 1 (Parameter restrictions) *We impose the following parameter restrictions throughout the analysis:*

It is worth clarifying what exactly is implied by the assumption that $\alpha_c < \alpha_p$. The α_j parameter denotes the likelihood of action j succeeding, *conditional on* the leader having private information suggesting that it might succeed. The total probability that action j can yield a successful outcome is given by $\tau_j \alpha_j = \sum_{\omega_j} Pr(y = 1|a_j = 1, \omega_j) Pr(\omega_j)$. Assuming $\alpha_c < \alpha_p$ can still be consistent with covert action being more effective on average—that is, if $\alpha_c < \alpha_p < \frac{\tau_c}{\tau_p} \alpha_c$.

This assumption can be relaxed (i.e. allowing for $\alpha_c > \alpha_p$) without altering the core result regarding the cover-up mechanism. Likewise, we could relax point (vi), which would (under some conditions) imply that the unscrupulous leader optimally takes both public and covert action when both are feasible (that is, plays $\sigma^0(1, 1) = (\sigma_c > 0, \sigma_p > 0)$). Either of these changes would simply require more equilibrium cases to be solved for, and would distract from the main substantive points of the model.

Assumption 2 (Markovian strategies) *Let $t = (\theta, \omega)$. Restrict attention to equilibria in which, if $E[U_L(a) - U_L(a')|t] = E[U_L(a) - U_L(a')|t']$ for some $t \neq t'$, then $Pr(a|t) = Pr(a|t')$.*

Table A.1: Summary of Modeling Parameters

Parameter	Discussion of values
Strategic Environment Pre-play	
$\theta \in \{0, 1\}$	L 's privately-known type, scrupulous ($\theta = 1$) or unscrupulous ($\theta = 0$)
$\pi \in (0, 1)$	Prior prob. L is scrupulous, $Pr(\theta = 1) = \pi$.
$j \in \{c, p\}$	Subscript used to denote covert v. overt
$\omega_c, \omega_p \in \{0, 1\}$	Feasibility of each policy action: feasible ($\omega_j = 1$) or infeasible ($\omega_j = 0$), privately known to L
$\tau_c, \tau_p \in (0, 1)$	Prior prob. each action is feasible, $Pr(\omega_j = 1) = \tau_j$
Player actions	
$a_c, a_p \in \{0, 1\}$	L implements covert, overt policy (1 = yes, 0 = no).
$r \in \{0, 1\}$	1 = A rewards L ; 0 = A punishes L .
Strategic Environment following L 's action	
$y \in \{0, 1\} a, \omega$	policy success (1 = success).
$0 < \alpha_0 \leq \alpha_c \leq \alpha_p \leq \alpha_{pc} < 1$	$\Pr y = 1$ given L 's actions a their feasibility ω ; see Table 1
$z \in \{0, 1\} a_c, \lambda, \delta$	A observes direct evidence of L 's covert action (1 = A observes, 0 = A does not observe).
$\lambda \in (0, 1), \delta \in (0, 1 - \lambda)$	Determine the probability $z = 1$, depending on L 's actions. λ is the baseline risk of exposure given $a_c = 1$, $\lambda + \delta$ is amplified risk if L invites attention through overt action.
Payoff parameters	
$\beta > 0$	L 's reward from $r = 1$.
$k_p > 0$	L 's cost from implementing $a_p = 1$.
$k_c^\theta > 0$	L 's cost from implementing $a_c = 1$. Type-dependent. The unscrupulous type pays a lower cost.
$\bar{\mu} \in (0, \pi)$	A 's benefit from punishing L ($r = 0$).

This restriction follows from [Maskin and Tirole \(2001\)](#). It requires that strategies are conditioned only payoff-relevant information. Intuitively, if the leader has the same expected payoff from each of his actions under two signal realizations, we have no substantive reason to focus on equilibria that rely on him behaving differently under those two signal realizations.

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

- When public action is feasible, the leader only takes public action: $\sigma(\omega_c, 1) = (0, 1)$.
- When neither action is feasible, the leader takes neither: $\sigma(0, 0) = (0, 0)$.

A.1 Equilibrium Existence

Outline:

- Lemmas 1, 2, and 3 characterize behavior within any equilibrium satisfying Assumptions 1, 2, and 3.

	Assumption	Interpretation/Justification
(i)	$\kappa_c^{\theta=1} > 1 + \beta$	Ensures that the scrupulous leader never takes covert action
(ii)	$\alpha_c < \alpha_p$	Leader has better info re: effectiveness of public vs. covert action
(iii)	$\alpha_0 \beta < k_p$	Upper bounds on β : reputational concerns
(iv)	$\beta(1 - \alpha_p - \alpha_0) < \alpha_p - k_p - \alpha_0$	are not too great, relative to policy benefits and costs
(v)	$\beta(-\alpha_p + (1 - \lambda)(1 - \alpha_c)) < (\alpha_p - k_p) - (\alpha_c - k_c)$	
(vi)	$\alpha_{pc} \leq \min \left\{ \alpha_p + k_c, \frac{\alpha_p}{1 - \lambda - \delta} \right\}$	When feasible public action is being taken, the additional returns to covert action are not too great
(vii)	$\frac{k_p}{\beta} \geq -\delta\alpha_c + (2\alpha_c - 1)(1 - \lambda)$	Ensures that the leader is not punished following successful public action and no covert revelation (that is, in the information set $(a_p = 1, y = 0, z = 0)$)
(viii)	$0 < \bar{\mu} < \pi$	Audience does not punish leader given prior beliefs of leader's scruples.
(ix)	$\alpha_0 < \min\{\alpha_p - k_p, \alpha_c - k_c^0\}$	Either action is (independently) net beneficial (for $\theta = 0$) for policy reasons alone, when feasible

- Proposition 2 builds from these lemmas to show that under Assumption 1, an equilibrium satisfying Assumptions 2 and 3 can always be supported.

Lemma 1 *Any equilibrium satisfying Assumptions 1, 2, and 3 will feature the following strategy for the leader:*

		Public action feasibility	
		$\omega_p = 1$	$\omega_p = 0$
Covert action feasibility	$\omega_c = 1$	$\sigma(\omega) = (0, 1)$	$\sigma^1(\omega) = (0, 0)$ $\sigma^0(\omega) = (\hat{\sigma}_c, a_c \hat{\sigma}_p)$
	$\omega_c = 0$		$\sigma(\omega) = (0, 0)$

Table A.2: Leader's equilibrium strategy, in any equilibrium satisfying Assumptions 1, 2, and 3

Proof of Lemma 1: Here we want to show that:

- In any equilibrium satisfying Assumptions 1 and 2;

- If $\sigma(0, 0) = (0, 0)$ and $\sigma(\omega_c, 1) = (0, 1)$, then:

- (i) $\sigma^1(1, 0) = (0, 0)$; and
- (ii) $\sigma^0(1, 0) = (\hat{\sigma}_c, a_c \hat{\sigma}_p)$; or in other words, $Pr(a_p = 1 | a_c = 0, \omega = (1, 0)) = 0$.

(i). From the assumption that $k_c^{\theta=1} > 1 + \beta$, it follows that the scrupulous leader will never take covert action, as it is too costly to offset even the best possible policy and reputational benefits it could bring. It is assumed that the scrupulous leader plays $\sigma^1(0, 0) = (0, 0)$; and because he does not take covert action, the value of ω_c is payoff-irrelevant to his choice of a_p ; so by Assumption 2, he must play $\sigma^1(1, 0) = (0, 0)$.

(ii). Analogously to the previous point: It is assumed that the unscrupulous leader plays $\sigma^0(0, 0) = (0, 0)$; when he does not take covert action, the value of ω_c is payoff-irrelevant to his choice of a_p ; so by Assumption 2, he must play $\sigma_p^0(1, 0) = 0$ when $a_c = 0$.

■

Audience beliefs and strategies Let $\mu^h = Pr(\theta = 1 | h)$ denote the audience's posterior belief of the leader's scruples, and let $\hat{r}^h = Pr(r = 1 | h)$ denote the audience's reward/punishment strategy, in information set $h = (a_p, y, z)$. For notational convenience, let $q = \hat{r}^{010}$, $s = \hat{r}^{110}$, and $t = \hat{r}^{100}$.

Lemma 2 *In any equilibrium satisfying Assumptions 1, 2, and 3, the audience's posterior beliefs are determined by Bayes' Rule to satisfy:*

$$\begin{aligned}
\mu^{a_p, y, 1} &= 0 \quad \forall a_p, y \\
\mu^{000} &= \frac{\pi}{\pi + (1 - \pi) \left\{ 1 - \tau_c \hat{\sigma}_c \left[1 - (1 - \hat{\sigma}_p) \left(\frac{1 - \alpha_c}{1 - \alpha_0} \right) (1 - \lambda) \right] \right\}} \geq \pi \\
\mu^{010} &= \frac{\pi}{\pi + (1 - \pi) \left\{ 1 + \tau_c \hat{\sigma}_c \left[(1 - \hat{\sigma}_p) \frac{\alpha_c}{\alpha_0} (1 - \lambda) - 1 \right] \right\}} \leq \mu^{000} \\
\mu^{1y0} &= \frac{\pi}{\pi + (1 - \pi) \left\{ 1 + \left(\frac{1 - \tau_p}{\tau_p} \right) \tau_c \hat{\sigma}_c \hat{\sigma}_p \left(\frac{\hat{y}_c}{\hat{y}_p} \right) (1 - \lambda - \delta) \right\}} \tag{3}
\end{aligned}$$

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

Consequently, the audience's equilibrium strategy will satisfy:

$$\begin{aligned}
\hat{r}^{a_p, y, 1} &= 0 \\
\hat{r}^{000} &= 1
\end{aligned}$$

Proof of Lemma 2:

The beliefs follow directly from Bayes' Rule, given the leader strategy specified in Table A.2.

From (1), we can see that $E[U_A(r = 1) | a_p, y, z] = \mu^{a_p, y, z}$, and $E[U_A(r = 0) | a_p, y, z] = \bar{\mu}$. The audience strategy stated in the lemma follows from the fact that $\mu^{a_p, y, 1} = 0 < \bar{\mu} < \pi \leq \mu^{000}$.

■

Define $\Delta_p(\hat{r})$ as the leader's payoff from playing $a_p = 1$ over $a_p = 0$ given $\omega = (1, 0)$ and $a_c = 1$, and given the audience's strategy $\hat{r} = (\hat{r}^{010}, \hat{r}^{110}, \hat{r}^{100}) = (q, s, t)$. That is:

$$\begin{aligned}\Delta_p &= E[U_L(a_p = 1) - U_L(a_p = 0)|\omega = (1, 0), a_c = 1] \\ &= -k_p + \beta \{ \alpha_c [(1 - \lambda - \delta)s - (1 - \lambda)q] + (1 - \alpha_c) [(1 - \lambda - \delta)t - (1 - \lambda)] \}\end{aligned}\quad (4)$$

Lemma 3 Denote the equilibrium strategy $\hat{\sigma}_p = Pr(a_p = 1|\omega = (1, 0), a_c = 1, \theta = 0)$. In any equilibrium satisfying Assumptions 1, 2, and 3:

- $\hat{\sigma}_p < 1$.
- If $\hat{\sigma}_p > 0$, then $\Delta_p = 0$.
- If $\hat{\sigma}_p = 0$, then $t = 1$.
- If $\hat{\sigma}_p > 0$, then $t \geq q$.
- $s = 1$.

Proof of Lemma 3: We will consider each point of the lemma in turn.

- $\hat{\sigma}_p < 1$.

Suppose $\hat{\sigma}_p = 1$. Then $\mu^{010} \geq \pi > \bar{\mu}$, so $\hat{r}^{010} = 1$. But $\Delta_p(q = 1, s, t) < 0$, meaning that L has a profitable deviation to $a_p = 0$ given $\omega = (1, 0), a_c = 1$, contradicting $\hat{\sigma}_p > 0$. So it must be that $\hat{\sigma}_p < 1$.

- If $\hat{\sigma}_p > 0$, then $\Delta_p = 0$.

$\hat{\sigma}_p > 0$ implies that $E[U_L(a_p = 1)|\omega = (1, 0), a_c = 1] \geq E[U_L(a_p = 0)|\omega = (1, 0), a_c = 1]$. If this inequality were strict, then L would strictly prefer $a_p = 1$, which would mean $\hat{\sigma}_p = 1$, which was the previous point showed cannot hold. So it must be that $E[U_L(a_p = 1)|\omega = (1, 0), a_c = 1] = E[U_L(a_p = 0)|\omega = (1, 0), a_c = 1]$, that is, $\Delta_p = 0$.

- If $\hat{\sigma}_p = 0$, then $t = 1$.

Follows directly from the fact that $\mu^{1,0,0} = \pi > \bar{\mu}$ when $\hat{\sigma}_p = 0$, as per Lemma 2.

- If $\hat{\sigma}_p > 0$, then $t \geq q$.

Observe that $\Delta_p = 0$ rearranges to

$$t = \frac{\frac{k_p}{\beta} - \alpha_c(1 - \lambda - \delta)s + \alpha_c(1 - \lambda)q + (1 - \alpha_c)(1 - \lambda)}{(1 - \alpha_c)(1 - \lambda - \delta)}$$

Rearranging further, we find that this is $\geq q$ whenever $\frac{k_p}{\beta} \geq -\delta\alpha_c + (2\alpha_c - 1)(1 - \lambda)$, as stipulated in point (vii) of Assumption 1.

- $s = 1$

First observe that $\mu^{110} > \mu^{100}$ when $\hat{\sigma}_p > 0$ (as per Lemma 2). For the audience to play a fully mixed strategy in information set h requires that $\mu^h = \bar{\mu}$, so the audience can fully mix for only one of $h = (1, 1, 0)$ or $h = (0, 1, 0)$. So the ordering of $\mu^{110} > \mu^{100}$ implies $s \geq t$, with at most one of $s \in (0, 1)$ or $t \in (0, 1)$.

Now, suppose $s < 1$. This implies that $t = 0$, and by the previous points, that $q = 0$, and that $\hat{\sigma}_p > 0$. Plugging these values into $\Delta_p = 0$ (which is implied by $\hat{\sigma}_p > 0$), we have

$$\frac{k_p}{\beta} = \alpha_c(1 - \lambda - \delta)s - (1 - \alpha_c)(1 - \lambda)$$

which, given $s < 1$, contradicts the assumption that $\frac{k_p}{\beta} \geq -\delta\alpha_c + (2\alpha_c - 1)(1 - \lambda)$. Therefore it must be that $s = 1$.

■.

Proposition 2 *Under Assumption 1, an equilibrium satisfying the conditions of Assumptions 2 and 3 can always be supported.*

Proof of Proposition 2:

We will prove the proposition by showing that the strategies specified in Lemma 1 are supported in equilibrium under Assumption 1. (It is self-evident that these strategies satisfy Assumptions 2 and 3.)

Thus to prove equilibrium existence, we need to show the following incentive-compatibility conditions:

- (i) $\sigma^\theta(0, 0) = (0, 0)$ for $\theta = 0, 1$ (which implies $\sigma^1(1, 0) = (0, 0)$)
- (ii) $\sigma^\theta(0, 1) = (0, 1)$ for $\theta = 0, 1$
- (iii) $\sigma_p^0(1, 0) = a_c \hat{\sigma}_p$ for some $\hat{\sigma}_p \in [0, 1]$ (that is, $Pr(a_p = 1 | a_c = 0, \omega = (1, 0)) = 0$)
- (iv) $\sigma^\theta(1, 1) = (0, 1)$ for $\theta = 0, 1$

We will consider each in turn.

(i) *IC(0, 0)*: $\sigma^\theta(0, 0) = (0, 0)$ for $\theta = 0, 1$.

In state $\omega = (0, 0)$, it is clear that neither leader would benefit from deviating to $a_c = 1$: doing so is intrinsically costly, carries no policy benefit, and yields a worse expected reputational payoff (weakly worse when $\lambda = 0$, strictly worse otherwise). The only deviation to consider is from $a_p = 0$ to $a_p = 1$. The leader's expected payoffs from each action are:

$$\begin{aligned} E[U_L(0, 0) | \omega = (0, 0)] &= \alpha_0 + \beta(\alpha_0 q + (1 - \alpha_0)) \\ E[U_L(0, 1) | \omega = (0, 0)] &= \alpha_0 - k_p + \beta(\alpha_0 + (1 - \alpha_0)t) \end{aligned}$$

The non-deviation condition requires

$$\beta[\alpha_0(1 - q) - (1 - \alpha_0)(1 - t)] \leq k_p$$

which is satisfied by $\beta\alpha_0 \leq k_p$, as per point (iii) of Assumption 1.

Note that $IC^1(0, 0)$ implies $IC^1(1, 0)$: given that the scrupulous leader does not take covert action, his best-response in state $\omega = (0, 0)$ coincides with his best-response in state $\omega = (1, 0)$.

(ii) $IC(0, 1) : \sigma^\theta(0, 1) = (0, 1)$ for $\theta = 0, 1$.

In state $\omega = (0, 1)$, we can again see that neither leader type will deviate to $a_c = 1$, for the same reasons given above. The deviation we need to consider is from $a_p = 1$ to $a_p = 0$. The leader's expected payoffs from each action are as follows:

$$\begin{aligned} E[U_L(0, 0)|\omega = (0, 1)] &= \alpha_0 + \beta(\alpha_0 q + (1 - \alpha_0)) \\ E[U_L(0, 1)|\omega = (0, 1)] &= \alpha_p - k_p + \beta(\alpha_p + (1 - \alpha_p)t) \end{aligned}$$

The non-deviation condition requires

$$\beta [\alpha_0 q + (1 - \alpha_0) - \alpha_p - (1 - \alpha_p)t] \leq \alpha_p - k_p - \alpha_0$$

which is satisfied by point (iv) of Assumption 1. (Recall that $q \leq t$ in equilibrium.)

(iii) $IC^0(1, 0) : \sigma_p^0(1, 0) = a_c \hat{\sigma}_p$ for some $\hat{\sigma}_p \in [0, 1]$.

This follows directly from $IC^0(0, 0)$: given $a_c = 0$, the payoff from $a_p = 0$ vs. $a_p = 1$ is unaffected by ω_c . Thus if $\sigma_p^0(0, 0) = 0$ is supported in equilibrium, then so is $\sigma_p^0(1, 0) = 0$ given $a_c = 0$.

(iv) $IC(1, 1) : \sigma^\theta(1, 1) = (0, 1)$ for $\theta = 0, 1$.

Any deviation away from $\sigma(1, 1) = (0, 1)$ is (weakly) more appealing for the unscrupulous than the scrupulous leader, so we will consider the unscrupulous leader's IC condition. There are three possible deviations: to $a = (0, 0)$, to $a = (1, 0)$, and to $a = (1, 1)$. The non-deviation to $a = (0, 0)$ was shown in $IC(0, 1)$ above.

The leader's expected payoffs from each action are as follows:

$$\begin{aligned} E[U_L(0, 1)|\omega = (1, 1)] &= \alpha_p - k_p + \beta(\alpha_p s + (1 - \alpha_p)t) \\ E[U_L(1, 0)|\omega = (1, 1)] &= \alpha_c - k_c + \beta(1 - \lambda)(\alpha_c q + (1 - \alpha_c)) \\ E[U_L(1, 1)|\omega = (1, 1)] &= \alpha_{pc} - k_p - k_c + \beta(1 - \lambda - \delta)(\alpha_{pc} s + (1 - \alpha_{pc})t) \end{aligned}$$

L does not want to deviate from $a = (0, 1)$ to $a = (1, 0)$ if

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

which is satisfied by point (v) of Assumption 1.

L does not want to deviate from $a = (0, 1)$ to $a = (1, 1)$ if

$$\alpha_{pc} - \alpha_p - k_c \leq \beta [(\alpha_p - (1 - \lambda - \delta)\alpha_{pc})s + ((1 - \alpha_p) - (1 - \lambda - \delta)(1 - \alpha_{pc}))t]$$

which is satisfied by point (vi) of Assumption 1.

This exhausts all incentive-compatibility conditions for the leader; thus an equilibrium satisfying the conditions of Assumptions 2 and 3 can always be supported under Assumption 1.

■

A.2 Comparative Statics

Outline:

- Proposition 3 states the conditions under which the leader will take covert action, in the equilibrium characterized in Proposition 2.
- Proposition 4 states the conditions under which the leader will employ a cover story, in the equilibrium characterized in Proposition 2.
- Proposition 1 shows that the class of equilibrium that is supported is a function of two threshold values in λ .
- Corollary 2 formalizes the counterintuitive result depicted in the main text, showing that the probability of using cover stories can increase in the additional scrutiny that they invite.

Remark The remainder of the appendix will focus on behavior by the unscrupulous leader (suppressing $\theta = 0$ from the notation), in an equilibrium satisfying Assumptions 1, 2, and 3. (The scrupulous leader’s behavior is as specified in Lemma 1, and—unlike the unscrupulous leader’s behavior—does not vary as a function of exogenous parameters.)

Lemma 4 Let $\hat{\lambda} := 1 - \frac{\alpha_0(\psi-1+\tau_c)}{\alpha_c\tau_c}$, , where $\psi := \frac{\pi(1-\bar{\mu})}{(1-\pi)\bar{\mu}}$. If $\lambda > \hat{\lambda}$, then $\mu^{010} > \bar{\mu}$ and thus $q = 1$ for any $\hat{\sigma}_c, \hat{\sigma}_p$.

Proof of Lemma 4:

From Lemma 2, we can see that μ^{010} is minimized when $\hat{\sigma}_c = 1$ and $\hat{\sigma}_p = 0$; plugging in these values, we find that $\mu^{010} > \bar{\mu}$ whenever $\lambda > \hat{\lambda}$. ■

The main takeaway of Lemma 4 is that, when transparency is sufficiently high, “no news is good news”: that is, the absence of revelation of covert action means that it is highly unlikely that covert action took place. Under this condition, upon observing a successful policy outcome without public action, the audience is willing to give the leader the benefit of the doubt, rather than punishing him out of suspicion of unobserved covert action.

Lemma 5 outlines the leader’s best-response covert action strategy as a function of the audience’s punishment strategy, showing that the leader’s probability of using covert action is (weakly) decreasing in the probability that he gets punished in the $h = (a_p = 0, y = 1, z = 0)$ information set.

Lemma 5 Define Δ_c as:

$$\begin{aligned} \Delta_c(q) &= E[U_L(a_c = 1)|\omega = (1, 0)] - E[U_L(a_c = 0)|\omega = (1, 0)] \\ &= \alpha_c - \alpha_0 - k_c - \beta \{ \lambda + (1 - q) [(1 - \lambda)\alpha_c - \alpha_0] \} \end{aligned} \quad (5)$$

L ’s best-response covert action strategy in state $\omega = (1, 0)$, and given audience strategy \hat{r} , satisfies

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

where $\Delta_c(q)$ is given by (5). If $\lambda < \hat{\lambda}$, then $\Delta_c(q)$ is increasing in q .

Proof of Lemma 5:

We showed above that $\hat{\sigma}_p < 1$. 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 $a_p = 0$.

(5) follows directly from the leader's payoff function (and from the fact that $\hat{r}^{000} = 1$, as per Lemma 2). Then the claim that $\Delta_c(q)$ is increasing in q follows simply from the fact that $\lambda < \hat{\lambda} \implies (1 - \lambda)\alpha_c - \alpha_0 > 0$. ■

Proposition 3 *The leader takes covert action with positive probability, $\hat{\sigma}_c > 0$, 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 3:

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

- Suppose $\Delta_c(1) < 0$. Either:

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

Second, to prove the “if”, $\hat{\sigma}_c > 0 \iff \Delta_c(1) > 0$:

- Suppose $\Delta_c(1) > 0$. If $\hat{\sigma}_c = 0$, 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 $\hat{\sigma}_c = 0$. So $\Delta_c(1) > 0 \implies \hat{\sigma}_c > 0$.

■

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

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

Lemma 6 *Let $\tilde{q} := \begin{cases} \frac{-\Delta_c(0)}{\Delta_c(1) - \Delta_c(0)}, & \lambda \leq \hat{\lambda} \\ 1, & \lambda > \hat{\lambda} \end{cases}$, and let $\hat{q} = \max\{0, \min\{\tilde{q}, 1\}\}$. If the leader never uses cover stories (that is, if $\hat{\sigma}_p = 0$), then the audience's equilibrium strategy satisfies $q = \hat{q}$.*

Proof of Lemma 6:

Observe that \tilde{q} is the value of q that would make L indifferent between taking covert action or not, in the $\omega = (1, 0)$ state, as per Lemma 5.

- If $\tilde{q} \geq 1$, this means that the leader does not take covert action under any possible value of q (that is, any $q \in [0, 1]$); thus in equilibrium, $\hat{\sigma}_c = 0$ and $q = 1$.
- If $\tilde{q} \leq 0$, this means that the leader will always take covert action under any possible value of q ; thus, in equilibrium, $\hat{\sigma}_c = 1$; and because $\lambda \leq \hat{\lambda}$ and $\hat{\sigma}_p = 0$, it follows that $\mu^{010} < \bar{\mu}$, so $q = 0$.
- If $\tilde{q} \in (0, 1)$, this means that $\Delta_c(0) < 0 < \Delta_c(1)$, which means that the leader must play a fully mixed covert action strategy; in order to maintain the leader's indifference to support mixing, the audience must play $q = \tilde{q}$.

■

Proposition 4 *The leader uses a cover story with positive probability, $\hat{\sigma}_c \hat{\sigma}_p > 0$, 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 4:

Intuitively, the proof proceeds as follows: Suppose that the leader is not using cover stories (which, as per Lemmas 3 and 6, implies that $q = \hat{q}$ and $s = t = 1$); and show the conditions under which he would deviate to using cover stories.

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

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

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

- (i). Suppose $\hat{\sigma}_c = 0$ and show $\Delta_p(q = \hat{q}, s = t = 1) \leq 0$.
 $\hat{\sigma}_c = 0 \implies s = t = 1$. From Proposition 3, $\hat{\sigma}_c = 0 \implies \Delta_c(1) \leq 0$. If $\lambda \geq \hat{\lambda}$ then $\tilde{q} = 1$; otherwise $\Delta_c(0) < \Delta_c(1)$; in either case we have $\tilde{q} \geq 1$, which $\implies \hat{q} = 1$. So $\Delta_p(q = \hat{q}, s, t) = \Delta_p(q = 1, s = t = 1) < 0$.
- (ii). Suppose $\hat{\sigma}_c > 0$ and $\hat{\sigma}_p = 0$, and show $\Delta_p(q = \hat{q}, s = t = 1) \leq 0$.
 - If $\hat{q} = 1$, then $\Delta_p(\hat{q}, s, t) < 0$.
 - $\hat{\sigma}_p = 0 \implies s = t = 1$.
 - From (4) we know that $\hat{\sigma}_p = 0 \implies \Delta_p(q, s = t = 1) \leq 0$. Left to prove is that $q = \hat{q}$.
 - If $\hat{q} < 1$, then either:
 - * $\hat{\sigma}_c = 1$, which (along with $\hat{\sigma}_p = 0$ 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:
 - * $\hat{\sigma}_c \in (0, 1)$, meaning $\Delta_c(q) = 0$, so $q = \tilde{q} = \hat{q}$.

■

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

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

From Proposition 3, we have that Class I (and only Class I) equilibria are supported if and only if $\Delta_c(1) < 0$; and from Proposition 4, we have that Class III (and only Class III) equilibria are supported if and only if $\Delta_p(s = 1, t = 1, q = \hat{q}) > 0$. It follows that Class II (and only Class II) equilibria are supported if and only if both $\Delta_c(1) > 0$ and $\Delta_p(s = 1, t = 1, q = \hat{q}) < 0$.

To prove Proposition 1, we want to show that the conditions separating these equilibria correspond to thresholds in λ .

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

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

Differentiating with respect to λ gives

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

When $\lambda > \hat{\lambda}$, $\hat{q} = 1$ and thus $\Delta_p(\hat{q} = 1, s = 1, t = 1) < 0$. For $\lambda \leq \hat{\lambda}$: \tilde{q} is continuous and increasing in λ ; \hat{q} is continuous and (weakly) increasing in \tilde{q} ; and $\Delta_p(\hat{q}, s = 1, t = 1)$ is continuous and decreasing in \hat{q} . Thus either $\Delta_p(\hat{q}, s = 1, t = 1) < 0$ for all λ , or there exists a unique λ' such that $\Delta_p(\hat{q}, s = 1, t = 1) > 0$ if and only if $\lambda < \lambda'$.

■

Corollary 2 *Under Assumptions 1 and 3, if $\lambda \leq \hat{\lambda}$ and*

$$(\psi - 1)(X + W) < \tau_c W(X - 1), \quad \text{where} \quad X = \frac{\alpha_c}{\alpha_0}(1 - \lambda) \quad \text{and} \quad W = \left(\frac{1 - \tau_p}{\tau_p} \right) \left(\frac{1 - \alpha_c}{1 - \alpha_p} \right) (1 - \lambda - \delta)$$

then there exists an equilibrium in which $\frac{d\hat{\sigma}_p}{d\delta} > 0$: that is, the probability of the leader using a cover story (conditional on $a_c = 1$) is increasing in the additional scrutiny that the cover story invites.

Proof of Corollary 2: Let $\psi = \frac{\pi(1 - \bar{\mu})}{(1 - \pi)\bar{\mu}}$. Note that $\psi > 1$, and that $\lambda \leq \hat{\lambda} \implies X > 1$. We can rewrite

$$\mu^{100} = \frac{\pi}{\pi + (1 - \pi)(1 + \hat{\sigma}_c \hat{\sigma}_p \tau_c W)} \quad \text{and} \quad \mu^{010} = \frac{\pi}{\pi + (1 - \pi)(1 + \tau_c \hat{\sigma}_c [(1 - \hat{\sigma}_p)X - 1]}$$

Consider an equilibrium in which the audience's strategy satisfies $t \in (0, 1)$ and $q \in (0, 1)$. In this equilibrium, it must be the case that $\mu^{100} = \mu^{010} = \bar{\mu}$, which rearranges to

$$\psi - 1 = W\tau_c \hat{\sigma}_c \hat{\sigma}_p \quad \text{and} \quad \psi - 1 = \tau_c \hat{\sigma}_c ((1 - \hat{\sigma}_p)X - 1)$$

We want to show, first, that this equilibrium exists; and second, that in this equilibrium, $\hat{\sigma}_p$ is increasing in δ .

To show that $\hat{\sigma}_p$ is increasing in δ , we simply set the RHS of two expressions above equal to each other and rearrange, giving

$$\hat{\sigma}_p = \frac{X - 1}{X + W}$$

which we can see is increasing in δ .

To prove equilibrium existence, we want to show that, when $\hat{\sigma}_p = \frac{X-1}{X+W}$, there exists a $\hat{\sigma}_c$ that satisfies $\mu^{100} = \mu^{010} = \bar{\mu}$. Plugging in the value of $\hat{\sigma}_p$ and rearranging, we have

$$\hat{\sigma}_c = \frac{(\psi - 1)(X + W)}{\tau_c W (X - 1)}$$

which is ≤ 1 when the first condition stated in the corollary holds.

■