Doubling Down: The Dangers of Disclosing Secret Actions

Jacob Otto* William Spaniel†

September 5, 2019

Abstract

When an actor catches a state taking an objectionable secret action, it faces a dilemma. Exposing the action could force uncommitted states to terminate the behavior to save face. But it could also provoke committed states to escalate the activity now that others are aware of the infraction. We develop a model that captures this fundamental tradeoff. Three main results emerge. First, the state and its opponent may engage in a form of collusion—opponents do not expose committed states despite their distaste for the behavior. Second, when faced with uncertainty, the opponent may mistakenly expose a committed type and induce escalation, leading the opponent to have *ex post* regret. Finally, as the strength of secret action increases, states may engage in it less often. This counterintuitive result is a consequence of the opponent's greater willingness to expose, which deters less committed types from bluffing.

*Graduate Student, Department of Political Science, University of Pittsburgh.

†Assistant Professor, Department of Political Science, University of Pittsburgh. (williamspaniel@gmail.com, http://williamspaniel.com).

1 Introduction

Early in the Syrian Civil War, Iran provided aid in the form of funds and intelligence to the embattled Assad regime in Syria (Tisdall, 2012). In 2012, multiple international news agencies reported secret Iranian support for the Syrian government, including arms shipments seized by the Turkish government (Al Jazeera, 2012). The U.N. additionally published a report detailing Iranian arms shipments to the Syrian regime despite a weapons exports ban (Al Jazeera, 2012).

Western intelligence soon discovered that Iranian involvement in Syria went beyond the UN report, to include the creation and backing of Shia militia groups (Al Arabiya, 2012). Soon thereafter, Iran deployed 4000 soldiers to Syria to assist government forces (Fisk, 2013). Again, US officials condemned increased Iranian involvement, alleging that the Quds force created "coordinated attacks" and "trained militias" against rebel forces (Filkins, 2013). By 2014, Iranian presence had increased to 7000 armed forces (Dagher and Fitch, 2015). The depth of Iranian involvement only escalated from there. In 2017, Iran used ballistic missile from inside Iranian territory to attack a target in Syria (Tehran Times, 2017). Iran has even tried to expand the war with rocket attacks against Israel from Quds forces in Syria (Dagher and Fitch, 2015). Continual and increasing reporting of Iran's covert action in Syria has only led to a more openly defiant Iranian position in Syria.

Iran's deepening involvement in Syria highlights a tradeoff in revealing distasteful covert action. The guilty party may respond in two ways. From the exposer's perspective, the ideal case is that the state terminates the behavior to save face. But the decision could also backfire if trying to sidestep the reputational cost was the reason the state kept the action a secret in the first place. Once exposed, it becomes free to openly continue the policy and escalate the damage now that it no longer must cover its tracks. The fact that the exposer may not know how the state will respond complicates the situation and raises questions about how one should approach the situation. Recent research has examined escalation concerns in limited conflicts (Carson, 2016) and the effects of exposure of covert action by information and communications technology (Joseph and Poznansky, 2018). However, no study has investigated how the potential response and information problems influence the exposing actor's decision.

To address that, we develop a model of uncertainty and secret action. An aggressing

state begins by choosing whether to engage in a behavior that is not immediately observable to the broader public or international community.[1] If an interested party (the target state, an international institution, or a domestic political opponent) observes the action, it then chooses whether to expose the it. Exposing incurs a public reputation cost on the aggressor, which may then stop its behavior to save face or escalate it. The game then ends.

Central to our analysis, the interested party does not know how much the aggressor cares about the issues at stake. Consequently, it may not know how the aggressor will respond. As previewed above, exposure would make "incredible" aggressors desperate to save face and therefore terminate the action. But with the information out in the open and having nothing left to hide, "credible" aggressors would escalate the action.

Four central results emerge. First, if aggressor is likely a credible type, *all* types of aggressors choose the secret action. Incredible aggressors reason that, even if caught, the other side would not risk escalation. Credible aggressors are just too probable here, and the incredible types can exploit that fact.

Second, when credible aggressors are very unlikely, we observe more variation. Whenever the other side discovers the action, it would expose what has occurred. The least credible types therefore select out of secret action. The more credible types take the risk. Of these, middling types back down after exposure. However, the most credible types still escalate. For the other party, exposure remains worth the risk because those most credible types are rare.

Third, when credible aggressors are neither too common nor too uncommon, we observe the most variation. Here, the exposing with certainty would cause escalation relatively often, to the point that the actor no longer wishes to always publicize what it knows. However, it cannot always keep quiet either, otherwise the incredible types would take secret action too often. As a result, the actor sometimes chooses to expose and sometimes chooses not to. This has the interesting property of encouraging additional incredible types to take secret action—but not so many that the other party has a clear-cut decision to expose what it discovers.

This final case yields some intriguing comparative statics. Existing theories suggest

[1]The model later defines secret action through the structure of the game. However, it captures a wide variety of state behaviors, including: (1) clandestine actions, which no one is aware of when successful, (2) covert actions, which others can observe but cannot attribute the perpetrator, and (3) actions that are immediately obvious to some elites but not to the broader public.

that secret action becomes more likely as it becomes more effective and less observable to outsiders. Our model only corroborates those principles in the long run—i.e., as secret action becomes overwhelmingly strong and impossible to observe. In the more middling cases, the pivotal type of the aggressing state chooses its strategy to make the opponent unsure whether to expose or not. When secret action becomes stronger, the opponent becomes more inclined to expose because there is now more to gain by ending the action. As such, the slightly incredible types become more inclined to forgo secret action, which causes an overall decline in its probability. Meanwhile, increasing the probability of the opponent seeing the secret action means that the opponent need not expose the action as often conditional on that observation. We show that the net effect washes out, leading to no change in the probability that the aggressor takes secret action in the first place.

Our work is closest to that of Carson (2016). However, we diverge on two major assumptions. First, in Carson's (2016) framework, the target's leadership colludes with the aggressor's government. A hold up problem drives the mechanism—the leadership worries that its citizens will demand a response too extreme for the leadership's tastes, causing the leadership to keep quiet about the action. Here, the target considers remaining silent out of fear that the aggressor will escalate. Second, we allow incomplete information to play a role, which generates rich bluffing behaviors not previously observed. Properly accounting for those bluffs also means that we can speak to the aggressor's selection into secret action in the first place.

Methodologically, we bring a formal approach to the area of secret action, which is rare in this field.[2] Nevertheless, formal models have an advantage of enforcing accounting standards to ensure the validity of our argument. It is also useful because the inherent secrecy of the subject matter means that analysts rarely have complete data on the cases they care about. Potential for selection problems could also arise, and indeed our model confirms this. Furthermore, formalization helps clarify the scope of the argument. For example, Carson (2016) situates his work in the covert action sphere. We show that our central mechanisms also operates other realms of secrecy.

Nevertheless, understanding the substantive motivation behind the secret action we analyze is important. We therefore begin by microfounding the key assumptions

[2]The key exception (Spaniel and Poznansky, 2018) focuses on institutional design rather than the asymmetric information problem we consider.

underlying our model, and we illustrate the logic underlying our main theoretical results throughout.

2 Motivation

Formal models derive conclusions from the assumptions built into them. It is therefore worth spending a few moments to motivate the key incentives of secret action we wish to analyze.

A central tension in our model is why states may want to make an action secret. On one hand, public action is the more effective option if a state wants to achieve a policy goal. Secret action entails additional costs and risks not associated with over action. States have to pay concealment costs, such as sanitizing weapons, to hide their involvement in a covert action (Joseph and Poznansky, 2018). Additionally, if the covert action is revealed, democratic leaders may face domestic political costs for deceiving the public and lose the ability to appropriately justify the action (Joseph and Poznansky, 2018). Moreover, those other reasons aside, a state taking public action can accompany its subversion with the covert actions in its arsenal.

If taking public action is more effective, why do states go secret at all? The tradeoff is higher costs. States face potential backlash from engaging in subversive behavior. Sometimes, these are external. For example, Iran has seen increased sanctions for its support of Houthi rebels in the Yemeni Civil War (Greenwood, 2018). Russia has also been the target of sanctions from its support of Ukrainian separatists (Thompson, 2017). President Obama received cold receptions from allies when the Snowden leaks revealed sweeping U.S. spying operations (Edwards, 2015).

Other times, there may domestic ramifications. For instance, the Reagan administration faced backlash due to the exposure from the Iran-Contra affair. Exposure of Iranian involvement in Syria has led to protests in Iran and criticism from some Iranian politicians (Fathollah-Nejad, 2018). Revelations of UK intelligence and logistical support for covert US airstrikes in the current conflict in Yemen has led to calls for investigations into the extent of UK involvement (McVeigh, 2017). Taking public action guarantees that a state will suffer these backlash costs. Secret actions may be less effective, but they also give the state a chance at avoiding the negative publicity altogether.

Of course, going secret does not guarantee avoiding exposure. Another party may observe the choice. Such actors include an opposing state's intelligence agency, non-governmental organizations with watchdog arms, whistleblowers within the state taking the secret action, or investigative journalists. Observing the action gives that actor an important decision. It can expose the action or it can keep the state's secret.

Exposure may look attractive as a means to motivate the perpetrator to end it to save face. Continued defiance accrues costs in the form of trade sanctions and reduced foreign aid (Lebovic and Voeten, 2009). Domestic resistance groups may begin imposing political costs on the ruling regime (Dai, 2005). The Church committee, which investigated US covert action, greatly admonished the use of covert action by the intelligence community and previous administrations (Isenberg, 1989*b*). The Committees recommendations largely resulted in the restricted use of covert action until Reagan administration (Isenberg, 1989*b*). Socially motivated states may wish to back down simply to mitigate damage to their prestige and status (Finnemore and Sikkink, 1998). In the case of the USS *Pueblo* incident, where a US spy ship was captured by DPRK forces off the coast of North Korea, the US had to admit to the espionage, apologize, and assure DPRK leaders it would not happen again (Newton, 1992).

However, exposure is not a sure bet. Some states escalate following the exposure of their secret action. Recall how much of the benefit from secret action comes from skirting international backlash. Once exposed, some of that backlash cost is sunk. Stopping the action may reduce international and domestic hostility, but the government cannot return to the good graces it had when the world was in the dark. This has the perverse effect of making public action relatively *less* expensive than compared to when the state made the initial decision.

Indeed, we can see the perverse effect motivating states to escalate. During the Vietnam War the US engaged in Operation *MENU*, which involved the secret bombing of North Vietnamese Army positions in Cambodia (Lewis, 1976). Once revealed in 1969, the Nixon administration expanded the bombing operation under the code-name: *FREEDOM DEAL* which lasted for three years and encompassed a greater geographical area (Finney, 1973). Reporters noted that the Pentagon did not attempt to hide the escalation of the US bombing campaigns in Cambodia (Lewis, 1976) after the exposure of the previous secret option. By the 1970s, the bombings of Cambodia and Laos were considered by scholars to be an "open secret" (Cormac and Aldrich, 2018).

Why would an actor expose secret action if it only makes the problem worse? Existing explanations focus on the self-enrichment of the exposer (Terman, 2019). States who engage in it solidify their role as an upholder of international norms. NGOs benefit from increased prominence, and journalists gain direct value from the news publication. Whistleblowers may also wish to gain notoriety from their actions (Sagar, 2001, 2013).

In contrast, we explore a strategic problem that would-be exposers face. Such actors may be uncertain about how a state internalizes the costs associated with negative publicity, sanctions, and domestic resistance relative to the value of the benefits of subversive actions of all types. Conflict scholars describe this as *resolve*, and researchers have provided strong microfoundations for it as a source of asymmetric information. After all, the relative weighting of costs versus benefits is an internal characteristic of a leader (Wolford, 2007). Unless we can understand a leader's internal thought process, we cannot know her corresponding resolve. In turn, an exposer may suspect the state is unresolved and will back down once its secret action becomes public. But if the state is actually resolved, it seems that exposure could only exacerbate the problem.

Of course, understanding strategic dynamics under uncertainty is a challenging task. Building a model can help us obtain a better appreciation of the interaction, and so we develop one in the next section.

3 The Model

The game consists of two actors, A and B. We conceptualize A as a state weighing how to conduct some subversive policy. B is any actor that would prefer A take the least amount of subversive action and is in a position to potentially learn if A has conducted secret action. Rivals of state A, international organizations, people within A's government who disagree with the policy all fit within this scope.

Nature begins by drawing A's cost of exposure k from the interval $[\underline{k}, \overline{k}]$. As is standard in models of conflict, we conceptualize the uncertainty as over A's resolve—i.e., how much it cares about the outcomes versus the costs required to reach those outcomes. We only assume that the cumulative distribution function, called $F(k)$, is differentiable everywhere and strictly increasing on the interval.[3] State A observes the cost draw. In the ultimate iteration of the game, B does not, though we consider how the

[3]Thus, $f(k) > 0$ for all $k \in [\underline{k}, \overline{k}]$, where $f(k)$ is the probability density function.

complete information game unfolds as a stepping stone to the incomplete information solution.

After Nature draws A's type, A then chooses whether to take public subversive action, secret subversive action, or no action. Both the public and no action choices end the game. If A takes secret action, Nature reveals A's decision to B with probability p. Failure to reveal also ends the game. In contrast, following revelation, B decides whether to expose the action or not. Once more, not exposing ends the game. But if B does expose it, A has a final choice. It can escalate the action or quit.

Payoffs are as follows. If A begins with public action, it receives $V_P - k$, where V_P captures the subversive value of the action and the drawn k value captures the cost. As the previous section described, the k value implicitly incorporates A's resolve, or how much it cares that its flaunting of international norms is problematic.[4] B suffers $-V_P$.

If A begins by doing nothing, both parties receive 0. This reflects a status quo outcome.

If A takes secret action and Nature fails to reveal it to B or B does not expose it, A receives V_S, where $V_P > V_S > 0$. This captures the idea that public action is more powerful than secret action, if for no reason other than that A could conduct its secret actions alongside its public actions. B earns $-V_S$. Thus, the stronger the action A takes, the worse off B is, though A must also worry about the cost it pays for taking an action the public observes.[5]

If B exposes the secret action and A escalates, both players receive the same payoff as if A had led with public action straight away.[6] If B exposes and A quits in response, A earns $-\alpha k$ and B earns 0. Here, $\alpha \in (0, 1]$ captures how much A can save face by giving up its action. Higher values of α prevent course reversal from having much effect on A's standing.[7]

With the game now described, it is worth highlighting the strategic tension. From

[4]This is the same standardization trick that comes from the crisis bargaining literature, in which the cost of war implicitly covers a state's resolve (Fey and Ramsay, 2011, 157).

[5]We could also include the monetary costs of public and secret actions. However, adding these creates notational clutter without fundamentally altering our theoretical results. We therefore exclude them.

[6]We could reduce the value of public action to account for the delayed implementation. Like before, though, this does not fundamentally alter our results. We therefore again maintain equivalent payoffs.

[7]We could also allow A to maintain its now-exposed "secret" action, but the structure of the incentives makes this irrelevant. State A would receive $V_S - k$ for doing so. This is strictly worse than escalating to public action.

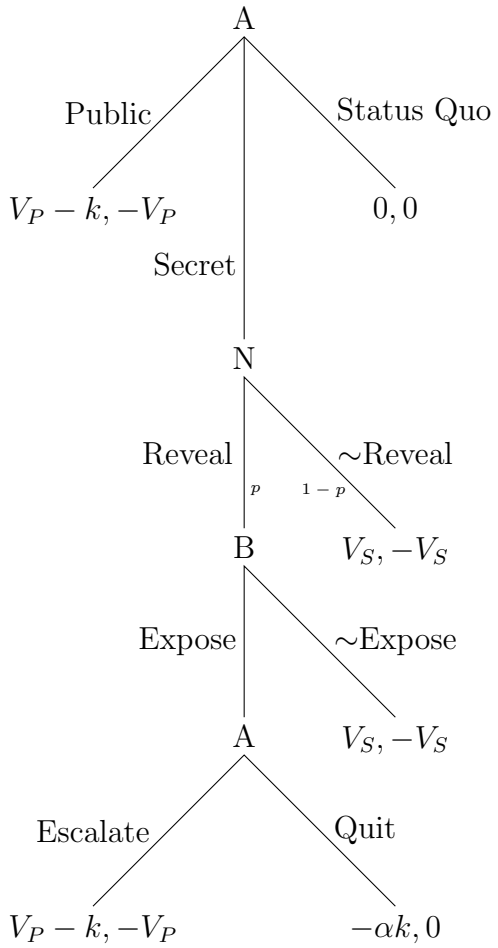

Figure 1: The extensive form of the complete information game.

a pure subversion perspective, A most prefers public action and least prefers no action. But it must also weigh its exposure cost. Going public incurs this cost directly. But secret action could still damage its reputation if Nature reveals it and then B publicizes it. So it may view secret action as more attractive than public action, though going secret is still a gamble.

Meanwhile, B's strategic dilemma is more involved. Exposing the secret action has no direct impact on its payoff.[8] However, it must consider the downstream consequences. A's subversion sinks a portion of A's exposure cost. Freed from that disincentive, A may escalate. Given that, B may wish to engage in a form of collusion, not publicize what it knows, and maintain A's course. On the other hand, A may wish to save face. In turn, B may wish to expose and thereby terminate A's action.

Incomplete information complicates both dilemmas. For B, it does not know how much A internalizes bad behavior. Thus, it may not anticipate how A will respond to exposure. Meanwhile, A's initiation influences B's belief. State A must therefore consider the signaling aspects of its decision.

Of course, not all parameter spaces feature these dilemmas. We therefore target our analysis by taking a single condition on the parameters. Namely, we focus on cases where $\underline{k} > V_P - V_S$. If $\underline{k}$ were lower than this value, any such type has a dominant strategy to take public action. As such, those types ignore the broader strategic considerations. Our discussion therefore focuses on the remaining cases.

3.1 Complete Information Equilibria

Although we mainly wish to analyze the incomplete information game, we begin with a complete information analysis. Doing so allows us to establish how the actors would behave if A's initial action revealed all relevant information about its type. This will generate better intuition for which types want to bluff and why, which builds toward the incomplete information case.

[8]This modeling choice ensures that our mechanism is distinct from existing explanations, which is a key modeling strategy of the "experimental" approach (Paine and Tyson, 2019). Recall that an actor may wish to expose secret action for personal benefit even if it exacerbates the action taken. If we included such a benefit in our model and also observed exposure backfiring, it may be unclear whether the personal benefit caused exposure to backfire or some other element of the strategic interaction. By excluding a personal benefit, we know that incomplete information drives the result. Adding a personal benefit does not yield additional theoretical insight. As such, we choose to analyze a model that gives the clearest elucidation of the mechanism.

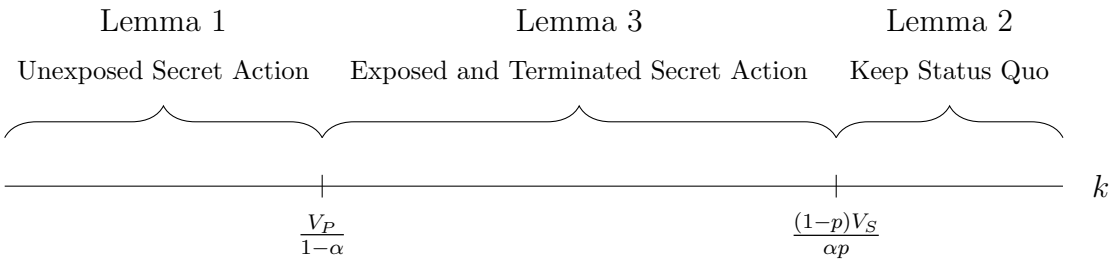

Figure 2: Parameter mapping for the complete information game.

As Figure 2 illustrates, the game falls into one of three cases based on the value of k:

3.1.1 When Costs Are Small

We begin with the case where A is not concerned about the punishment, as it leads to a straightforward conclusion:

Lemma 1. *Suppose the punishment cost is sufficiently low (i.e., $k < \frac{V_P}{1-\alpha}$). Then A engages in secret action. If revealed, B does not expose because—off the equilibrium path—A would escalate.*

When the potential punishment cost is low, A wants to run the risk of revelation to obtain the subversive benefits. But here, the costs are so low that it would escalate should B publicize the action. This concerns B. After all, it prefers A taking the more limited secret action than escalating to obtain the full effect of public action. As such, the threat of escalation deters B from publicizing, which in turn allows A to obtain its favorite outcome.

This parameter space formalizes a notion similar to Carson's (2016) state collusion. B observes that revealing the action would be counterproductive and only induce A to take more extreme action. State A benefits from saving on its exposure cost. B benefits from A's stinginess through the more moderate subversion strategy A chooses.

Lemma 1 contains another interesting implication. At first pass, B's ability to publicize A's decision would seem to deter A from taking secret action. Yet this is not always true. Note that exposure sinks $1 - \alpha$ portion of A's costs. That is, once B publicizes the action A can do nothing to recover that quantity. The effective discount

A receives at this point makes A more inclined to escalate than it would otherwise. In other words, exposure of A undermines B's ability to obtain the outcome it seeks.

To drive this point home further, consider A's utility for taking secret action, risking Nature informing B, being exposed, and having to escalate. It is possible that this quantity is less than if it just took no action at all. Working through the expectation, all the condition only requires is $p(V_P - k) + (1 - p)V_S < 0$, or $k < V_P + \frac{(1-p)V_S}{p}$. This must hold for the parameter space as the signal becomes perfectly informative, or p approaches 1. But A need not worry about this problem because it has a credible threat to escalate. In turn, B knows not to publicize, thereby turning A's apparently risky strategy into a safe bet.

In any case, Operation *RANCOUR* illustrates Lemma 1's central logic. During the North Yemen Civil War (1962–1972) multiple foreign powers intervened in the conflict both secretly and overtly, including the deployment of 70,000 Egyptian soldiers (Orkaby, 2017). While the Egyptians supported the republican forces, Saudi Arabia and Jordan provided military aid and the United Kingdom provided covert support for the royalists. The UK sought to hold on to the remnants of its empire, including the strategic port city of Aden in Yemen and continued access to oil reserves (Orkaby, 2017). Control of Aden became central to British foreign policy as hard-lining Conservatives within the government pushed for aggressively maintaining the UK's position in Yemen (McNamara, 2017). Given the anti-colonial domestic climate in the UK, external pressure from the US, and the deployment of Egyptian troops, the Foreign Office decided to engage in the covert support of royalist forces including the provision of weapons and mercenary support, according to a Top Secret memo (CoSC, 1964). Additionally, assassination of an Egyptian intelligence target was suggested but not implemented (Butler, 1964). However, overt military intervention was strongly considered in conjunction with British military troops in Aden, as declassified documents show (Macmillan, 1962).

Egyptian leaders, including Nasser, came to believe that there was covert British arming of royalist forces and threatened appropriate retaliation against the British colonial presence in Aden (McNamara, 2017). A declassified Foreign Office cable notes a meeting between the UK Ambassador in Cairo and a top Nasser advisor, who suspected British support for royalist forces, to which the UK ambassador denied the allegation (Beeley, 1962). British covert action continued throughout the conflict until 1968 and

plausible deniability was likely maintained through out (Cormac and Aldrich, 2018). In answering Parliamentary questions in 1964, the Prime Minister answered "Our policy towards the Yemen is one of non-intervention in the affairs of that country. It is not therefore our policy to supply arms to the Royalists in the Yemen and the Yemen government have not requested these or other forms of aid" (Hansard, 1964). Despite British covert action in South Arabia, Nasser appears to have not significantly attempted to reveal UK covert action.[9]

Fears of UK escalation to overt action by Egyptian strategists may have been warranted as the British strongly considered military intervention and at senior government officials at one point stated that exposure of the operation would not lead to a change in policy or a cessation of support for royalist tribes. Additionally, as the Egyptian military intervention drug on Nasser's position became more precarious as intelligence reports note abandoned military bases in Egypt and growing hostilities with Israel worsened the security position of the Egyptian homeland (Orkaby, 2017). South Arabia was only one hot spot in the struggle for influence in the Middle East between the UK and Egypt. UK escalation would only serve to worsen the Egypt's geostrategic position.

3.1.2 When Costs Are Large

Overall, Lemma 1 dealt with cases where A finds escalation acceptable because the cost of subversive action is manageable. The following lemma considers the opposite scenario:

Lemma 2. *Suppose the punishment cost is sufficiently high (i.e., $k > max\{\frac{V_P}{1-\alpha}, \frac{(1-p)V_S}{\alpha p}\}$). Then A maintains the status quo. Off the path, B would reveal secret action and A would quit.*

Again, the intuition is straightforward. When the punishment cost is high, A wishes to avoid incurring it as much as possible. Thus, it would quit the secret action if caught to safe face as much as possible. Anticipating that, B exposes the action to force A to

[9]Egyptian officials did at one point suggest UK support for republican forces in a UNSC brief but British declassified documents seem to be unconcerned about the mention. Additionally, some historical accounts note that Egyptian forces captured letters written by British mercenaries that were used in propaganda broadcasts in 1964. However, Nasser was potentially aware UK involvement in 1962 and did not publicize this information and there is some debate as to how covert the BMO mercenary operation was since it was run by a private company.

stand down. Realizing that will happen often enough, A avoids the mess entirely by maintaining the status quo from the start.

Obtaining Lemma 2 requires k to exceed two values for the following reasons. First, if A has a credible threat to escalate following revelation, then B would not want to reveal it. State A could in turn choose secret action without any fear, and this brings us to Lemma 1's case. But even if A would back down if revealed, it may still want to take a gamble. Indeed, if p is small, then A obtains the benefits of secret action a vast majority of the time, and the downside risk of exposure is almost nonexistent. As a result, and as Figure 2 shows, this parameter space disappears as p goes to 0. In contrast, as p goes to 1, this constraint becomes unimportant. Nature always reveals, and so A no longer has any real gamble. It simply looks at whether its threat to escalate is credible and chooses whether to take secret action accordingly.

Along the same lines, this parameter space disappears as α goes to 0. Under that extreme, A can almost entirely save face by withdrawing the secret action once exposed. In turn, A has no disincentive to try. If it works, A is happy. If it does not, then A can quietly withdraw without much consequence.

Obtaining empirical evidence of this mechanism is more challenging than before, because we would need to find evidence of secret actions that *never* happened. However, declassified records can provide some help here. One significant example of a covert operation that was considered at the highest levels but was denied due to the potential political costs that would incur if exposed was Operation *NORTHWOODS*. At the height of the Cold War and in the wake of the failed Bay of Pigs invasion, the Kennedy administration vigorously sought to overthrow the Castro regime in Cuba. In 1962, the Joint Staff proposed Operation *NORTHWOODS*. The idea reached the Chairman of the Joint Chiefs and the President before being discarded (OSD, 1962).

Operation *NORTHWOODS*'s strategic aim was to create a false-flag operation where the US could blame Cuba for the attacks and use them as a pretense eliminate the Castro regime (JCS, 1962). The plan called for shooting down US planes and hijacking US ships along with fabricating evidence implicating Cuba (JCS, 1962). The operation called for the targeting of US and UK citizens in coordinated terrorist attacks along the Atlantic coast (JCS, 1962). Cuban refugees would be targeted in the terrorist campaign using plastic explosives and refugee boats could be sunk en rout to Florida (JCS, 1962). Even the downing of a civil airliner was considered with passengers con-

sisting of "students off on a holiday" coming from a country where the "flight plan crosses over Cuba" (JCS, 1962). The base at Guantanamo Bay was a primary target and a Top Secret JCS memorandum suggests the US use "friendly" Cubans to stage attacks against the base, start riots near the main gate, fire mortars into the base, and sink ships in the harbor (JCS, 1962).

Documents on the plan only note that Kennedy declined to endorse the operation and removed the Chairman who proposed the plan (OSD, 1962). The costs of exposure were likely too great for Kennedy to accept as the direct targeting of US citizens by the US government would be an era/administration defining scandal and the domestic backlash would likely be severe. No amount of backtracking would allow the government to save any kind of face. Moreover, the threat of exposure was likely high coming from either potential whistleblowers, allied foreign governments, or the Castro regime publicizing its innocence. Given the existential threat the operation would pose to the Castro regime, Castro would have every incentive to publicize non-complicity in the attacks and expose US covert action.

3.1.3 When Costs Are Neither Small nor Large

A careful reader will note that, if $\frac{V_P}{1-\alpha} < \frac{(1-p)V_S}{\alpha p}$, Lemmas 1 and 2 do not cover a middle range of costs. Indeed, the constraints on Lemma 2 suggest that a more dynamic outcome may arise if the punishment cost is not so large. True to that, the following lemma shows that A may want to roll the dice under certain conditions:

Lemma 3. *Suppose the punishment cost falls in a middle region (i.e., $k \in (\frac{V_P}{1-\alpha}, \frac{(1-p)V_S}{\alpha p})$). Then A takes secret action and B exposes it if it has the opportunity to do so. If exposed, A quits.*

Here, the middling punishment cost induces A to take a gamble. It prefers to maintain the status quo than to engage in public action. But the likelihood that B will observe A's secret action is low. As such, it tries to get away with a little subversion but will back down when pressed. Note that this parameter space disappears if $\frac{V_0}{1-\alpha} > \frac{(1-p)V_S}{\alpha p}$. In that case, increasing k shifts the parameters directly from Lemma 1 to Lemma 2.

The Iran-Contra Affair serves as an illustration of this logic. Revelation of US arms sales to Iran during the Iran-Iraq War to fund CIA paramilitary activity in support of

the Contras in Nicaragua resulted in a national scandal. In 1984, Congressional disapproval of the US paramilitary activity in Nicaragua was dramatically rising, eventually culminating in the Boland amendment prohibiting any US agency from supporting anti-Sandinista efforts (Inouye and Hamilton, 1987). According to declassified documents, with full knowledge of Congress's position, the Reagan administration sought either financial support or material support from Brunei (DOS, 1986), Panama (North, 1986), and Honduras (McFarlane, 1985). Originally, the US sold weapons to Iran in exchange for the release of PLO hostages, including the controversial sale of 18 HAWK anti-aircraft missiles, using Israeli intelligence as an intermediary (CIA, 1985). Eventually, 120 HAWK missiles were sold, along with 500 TOW missiles and the US-provided intelligence on Iraqi military positions (Inouye and Hamilton, 1987). A congressional investigation revealed that the US made $16.1 million from arms sales to Iran and $3.8 million went to support the Contras (Inouye and Hamilton, 1987). Despite objections from his Secretary of State and Secretary of Defense, Reagan continued to push forward with weapons sales in exchange for the hostages as recorded in the diary of Secretary of Defense Caspar Weinberger (1985).

In 1986, NSC staffer Oliver North proposed in a Top Secret memo (the infamous "Diversion Memo") that the "residuals" of the Iranian arms deal be used to fund the Nicaraguan rebels (North, 1986). North directed the funds from the arms sales be directed to the Contras and to other covert action, according to a Congressional report (Inouye and Hamilton, 1987). The sale of weapons continued until the secret operation was exposed in 1986 after a senior official in Iran's Islamic Revolutionary Guard Corps leaked the secret action to a Lebanese magazine. However, there is evidence that the leak was arranged by NSC staff who had grown disenchanted with the operation and were concerned about the administration's disregard for the law. A few days after the publishing of the story, the Iranian government confirmed the story. The Reagan administration originally denied US arms sales to Iran before admitting to some sales. A declassified memorandum for record notes "the President said we did not do any trading with the enemy for our hostages" (Weinberger, 1985). The lack of information from the Administration regarding the scandal was largely due to a need to protect hostages and to protect lives in Iran and Nicaragua, according to the Secretary of Defense (Weinberger, 1986). Yet, fears of fallout after exposure persisted in the administration as Iran could stand to benefit from exposing the secret action. Even

the Secretary of Defense noted that the exposure of the story by the Iranians could compromise the administration by selectively exposing parts of the story (Weinberger, 1986).

The exposure of the US arms sales to an adversary created a significant political scandal for the Reagan administration and resulted in the termination of the secret arms sales to Iran. The costs of exposure of the secret weapons sales, including Congressional hearings, indictments, and significant public admonishment, created an environment which fostered the termination of the secret program. The political costs of exposure were likely multiplied by the fact that the secret action was illegal via the Boland amendment and the Executive did not inform Congress as required about the foreign arms sales. Reagan noted the illegality of his administration's secret action but remained resolved in administering the secret arms sales (Weinberger, 1985). Yet, upon the exposure of the secret action Reagan terminated the arms sales agreement as the costs of exposure was too high for the administration to bear.

3.2 Incomplete Information Equilibria

Taking stock of the complete information cases, the lowest cost types have little to hide. They want B to know they will escalate if exposed and convince B to accept a moderate amount of subversion. The other types are weaker, buckling upon exposure or not even trying. It would seem that they could benefit from the ability to bluff lower costs. We now explore that logic. And it is true—sometimes.

Consider the incomplete information model as originally formulated. The propositions below divide the parameter space by $F\left(\frac{V_P}{1-\alpha}\right)$. Recall that $F(k)$ is the cumulative distribution function defined on the interval $[\underline{k}, \overline{k}]$. Therefore, $F\left(\frac{V_P}{1-\alpha}\right)$ gives the portion of types that would escalate the secret action if exposed. As suggested a moment ago, this probability determines the higher cost types' ability to bluff. When large, upon observing secret action, B's initial inclination would seem to default to caution. In contrast, lower values would seem to inspire skepticism. We indeed see these reactions in the parameter spaces below.[10]

Throughout, we assume that $\underline{k} < \frac{V_P}{1-\alpha}$, which implies that $F\left(\frac{V_P}{1-\alpha}\right) > 0$. Making this

[10]We only describe A's initial action and B's exposure decision in the propositions below. For A's post-exposure strategies omitted from the propositions, see the corresponding lemmas. The appendix contains all proofs.

assumption allows us to focus on the most interesting case. If no such type existed, then A would always back down when exposed. The corresponding outcomes follow directly from Lemmas 2 and 3. Middle types would take secret action, and high cost types would not. Knowing that exposure ends the secret action, B publicizes it whenever Nature delivers that information. In contrast, as the previous paragraph outlined, having types with $k < \frac{V_P}{1-\alpha}$ means that B fears that its exposure could backfire, which gives rise to potential bluffs.

Indeed, we see such bluffs in our first parameter space:

Proposition 1. *Suppose low cost types are sufficiently likely (i.e., $F\left(\frac{V_P}{1-\alpha}\right) > \frac{V_S}{V_P}$). Then all types pool on secret action. B does not expose.*

Because low cost types are so frequent and B expects them to take secret action, it does not dare publicize any secret action revealed to it. The low cost types therefore choose secret action straight away. But anticipating B's tepid reaction, higher cost types also join in on the secret action. Although they would back down if challenged, they never have to worry about such a contingency. In fact, some types take secret action even though they have never would have in the complete information setting. Indeed, any type with k values falling under Lemma 2's parameters quit if B knows its high cost. But B does not know that information, and thus these types can bluff their way to a better outcome than before.

B's willingness to take chances constrains high cost types' ability to bluff. As Proposition 1 indicates, the portion of credible types cannot exceed the ratio $\frac{V_S}{V_P}$. Substantively, this cutpoint is how much worse public action is for B compared to secret action. When secret action causes almost as much damage as public action, V_S approaches V_P. In turn, B loses much of the downside risk of calling what it falsely sees as a bluff. It only suffers a marginal amount more when the credible types escalate. At the same time, B has a lot to gain. The non-credible types quit. Because secret action is close to public action in terms of strength, B's payoff increases by a wide margin. Consequently, under such conditions, bluffing only becomes possible for high cost types when those high cost types are unlikely.

This parameter space is the most strategically straightforward. Bluffs work, and high cost types do not fear any exposure. As such, Proposition 1 serves as a robustness check for the empirical implications that Lemma 1 generated. Even if we cannot know

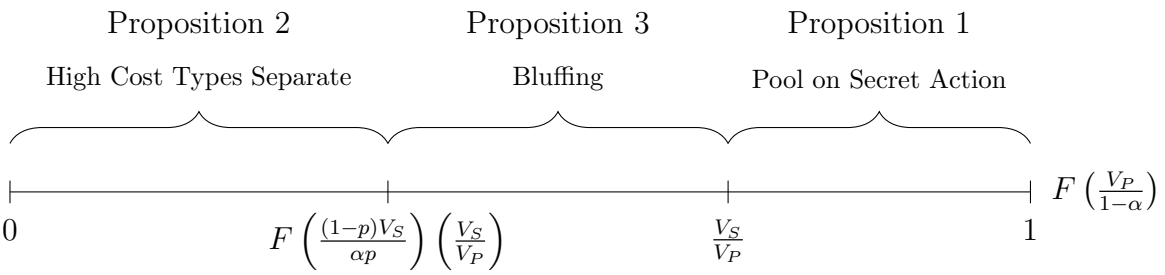

Figure 3: Parameter mapping for the incomplete information game.

with certainty that A's resolve was low, the *expectation* thereof predicts that A would take secret action and B would not expose. Cases like Operation *RANCOUR* follow from this.

In contrast, venturing away from Proposition 1's parameter space forces higher cost types to reassess their plan and leads to more strategically rich behavior. We now switch to situations where the most credible types are rare:

Proposition 2. *Suppose low cost types are sufficiently unlikely (i.e.,* $F\left(\frac{V_P}{1-\alpha}\right) < F\left(\frac{(1-p)V_S}{\alpha p}\right)\left(\frac{V_S}{V_P}\right)$*). Then all types with sufficiently low costs (i.e.,* $k < \frac{(1-p)V_S}{\alpha p}$*) take secret action, and the rest maintain the status quo. B exposes. The remaining highest cost types back down (i.e., those with* $k \in \left(\frac{V_P}{1-\alpha}, \frac{(1-p)V_S}{\alpha p}\right)$*), while the remaining lowest types escalate (i.e., those with* $k < \frac{V_P}{1-\alpha}$*).*

Here, middling types like those from Lemma 3 are pervasive. B knows they will test the waters. It therefore exposes the action, anticipating that those middling types will then fold. But B also knows that this strategy could backfire. Some portion of the time, A actually has low costs and escalates the subversion. Meanwhile, extremely high cost types know that B will challenge A, and so they sit out altogether.

Proposition 2 provides an empirical prediction that the complete information model does not. By the standard collusion logic, it is hard to explain why an actor would expose secret action if it only induces a worse response from the opponent. This model explains such an outcome as a result of uncertainty. B may suspect that exposing the action is the right decision, but bluffing middling types do not make that a sure bet. Rather, B exposes because it is the best move in expectation; it may have ex-post regret.

Note that we can only obtain this equilibrium outcome when Lemma 3's condition is filled. That is, some types must exist that wish to both take secret action and escalate if exposed, and some other types must exist that wish to take secret action but back down if exposed. The second group does not exist if, anticipating that B will expose anything it learns, taking secret action in the first place implies a type's preference to escalate later.

When the middle types do exist, they act like low cost types here by pooling with them on secret action. Nevertheless, their strategy is not particularly devious. They too would take secret action in a complete information setting, and thus Lemma 3 captures many of the incentives that Proposition 2 explores. It would be more surprising if even higher cost types opted for secret action despite B's temptation to call potential bluffs. We observe this in the final case, where the most credible types are neither too common nor too uncommon.

Before reaching that discussion, it will first help to further explore B's incentives. Suppose for the moment that all types with a cost less than some k value choose to take secret action. Then B's utility for exposing equals:

$$-\left(\frac{F\left(\frac{V_P}{1-\alpha}\right)}{F(k)}\right)V_P+\left(1-\frac{F\left(\frac{V_P}{1-\alpha}\right)}{F(k)}\right)0$$

Recalling that B earns $-V_S$ by maintaining silence, B is indifferent between exposing and not if:

$$-\left(\frac{F\left(\frac{V_P}{1-\alpha}\right)}{F(k)}\right)V_P+\left(1-\frac{F\left(\frac{V_P}{1-\alpha}\right)}{F(k)}\right)0=-V_S$$

$$\frac{F\left(\frac{V_P}{1-\alpha}\right)}{F(k)}=\frac{V_S}{V_P} \tag{1}$$

Let k^* be the unique solution to Line 1.[11] We are now ready for the final set of equilibrium strategies:

Proposition 3. *Suppose the likelihood of low cost types falls in a middle region (i.e., $F\left(\frac{V_P}{1-\alpha}\right)\in\left(F\left(\frac{(1-p)V_S}{\alpha p}\right)\left(\frac{V_S}{V_P}\right),\frac{V_S}{V_P}\right)$). Then all types with sufficiently low costs (i.e., $k<k^*$) take secret action, and the rest maintain the status quo. B sometimes exposes*

[11] We prove the existence and uniqueness of the solution in the appendix.

secret action and sometimes does not (i.e., it exposes with probability $\sigma^* \equiv \frac{V_S}{p(V_S+\alpha k^*)}$*). The remaining highest types back down (i.e., those with* $k \in \left(\frac{V_P}{1-\alpha}, k^*\right)$*) when exposed. The remaining lowest types (i.e., those with* $k < \frac{V_P}{1-\alpha}$*) escalate.*

Now the parameters impose deeper strategic problems for both players. For A, low cost types are frequent enough that B would not want to expose if the middle types from Figure 2 took secret action. But they are not so frequent that B would stay reticent if the highest cost types bluffed strength. And those high cost types would indeed want to, as they prefer taking secret action when B would not expose. Yet they cannot all take secret action, lest B begin wanting to call the potential bluff.

B is also in a conundrum. If it never exposes, high cost types will exploit it. But if it always reveals, those high cost types never enter the fray, and the option backfires on B too often.

The solution to each problem fixes the other. B must mix between exposing and not exposing. Doing so deters the highest types from taking secret action. After all, imagine k is so large that backing down produces an extremely negative payoff. Then even the slightest probability that B will expose it is enough to convince that type to not take the action in the first place. But the fact that B does not always expose here means that some of the higher cost types are willing to test their luck. These types fall in Lemma 2's range, where a pure exposure strategy would have otherwise deterred them. Because a greater portion of types that would back down now take secret action, B faces enough uncertainty that it is willing to mix between its strategies.

The more adventurous strategy from the higher cost types produces a new implication. Under both Lemma 3 and Proposition 2, the medium cost types back down when exposed in the game's equilibrium outcome. In both cases, A experiences *ex post* regret, but they prefer running that risk even if they know B would expose. The higher cost types also experience *ex post* regret here. But unlike before, if they knew B would expose, they would not have wanted to take the gamble in the first place. Rather, the uncertainty B induces by pursuing the mixed strategy convinces the higher cost types into attempting the secret action.

In the appendix, we calculate the equilibrium with a three-step process. First, we find the portion of the time exposing must backfire for B to be indifferent. Second, we use that probability to deduce a cutpoint on types k^*, for which all types less than k^* ought to take secret action. This range includes a portion of the highest types from

Figure 2. Thus, some types that would have otherwise declared defeat from the start now challenge B to call their bluff. Finally, we use this k^* to derive B's mixed strategy, such that the k^* type is indifferent. This ensures that all types greater than k^* have a pure preferences to maintain the status quo, while all types less than k^* have a pure preference to take secret action.

Like similar models of secret action (Spaniel and Poznansky, 2018), we invite readers to interpret B's mixed strategy through the purification theorem (Harsanyi, 1973). That is, rather than think of B literally mixing between its strategies, we could instead consider the class of neighboring games where A has small amounts of incomplete information over A's preference over the status quo. Types of B with marginal preferences for the status quo (or any other payoff) expose A as a pure strategy, hoping to acquire the very small additional benefit when A backs down. Meanwhile, types of B with marginal preference against the status quo do not expose as a pure strategy, very slightly wishing not to reach the status quo outcome. The purification theorem guarantees that the pure strategy of that game will have strategy and outcome probabilities converge to the mixed strategy equilibrium presented here as A's uncertainty goes to 0.

The central lesson from Proposition 3 is that actors may expose distasteful policies only for that decision to backfire. The shift from covert to overt military aid by the CIA in Afghanistan illustrates how Soviet exposure of US covert action resulted in a less preferable policy outcome. With the 1979 Soviet invasion of Afghanistan, President Jimmy Carter authorized a secret CIA program, Operation *Cyclone*, which provided support for the Afghan rebels, the *mujahideen* (Brzezinski, 1979). President Carter's national security advisor wrote to him explaining that the US now had the opportunity to give the Soviets their own "Vietnam War" (Gibbs, 2000).

One reason why the operation was initially covert was because of a hostile domestic climate in the US regarding executive overreach. The 1970s saw significant criticism and backlash pointed at intelligence agencies, namely manifested in the Church Committee report on curtailing intelligence activities (Isenberg, 1989*a*). However, covert action saw a resurgence under President Reagan. The Reagan administration expanded the program to include weapons and training, facilitated by the Pakistani ISI, in an effort to pressure the Soviets into withdrawal (Pach, 2006). Soviet intelligence likely picked up on American covert action and as early as 1980 reports from the Soviets detailed foreign

involvement in Afghanistan which was reported in the American press. One *New York Times* article from January 1980 was titled "Kremlin Steps Up Anti-U.S. Campaign, Charges CIA Is Training Afghan Rebels" (Fisher, 1980) as the Soviets sought to increase the pressure on the US. In a secret memorandum, Soviet intelligence assessed that the US was supplying arms to Afghan rebels through third-party states, such as Egypt and Saudi Arabia, to Pakistani intelligence to be smuggled across the border (Ustinov, 1980).

As the conflict continued the Reagan administration authorized the use of US-made Stingers, an anti-aircraft weapon that was uniquely identifiable as American and had a controlled proliferation (Carson, 2016; Pear, 1988). Since it was produced in 1983, the Stinger was considered cutting edge anti-air technology that would significantly hamper Soviet helicopter-borne counter-insurgency forces (Carson, 2016). The presence of Stingers on the battlefield signified an escalation in the conflict as now there was relatively undeniable evidence that the US was complicit in supporting the mujahideen, which came with possible significant ramifications. As Kuperman (1999) notes: "For the Central Intelligence Agency and especially its cautious Deputy Director John McMahon, directly traceable U.S. involvement raised the danger of public exposure and political scandal that could damage the agency, as had earlier CIA covert operations uncovered by the Pike and Church congressional committees in the 1970s". Yet, the Reagan administration eventually approved the Stingers. A Soviet intelligence report notes that that US likely delivered at least 600 Stingers to the Afghan rebels and over 250 US advisers were in Afghanistan training 100 rebels on Stinger use (GRU Memorandum, 1988).

The decision to escalate to overt action was likely due to domestic calls for more aid to rebels. In fact, it appears that the exposure of Operation Cyclone resulted in little blow back against the Reagan administration and Congressional support for the Afghan rebels continued as funding to the CIA program continued to increase until 1991 (Cogan, 1993). The chairman of defense appropriations, Charlie Wilson, was instrumental in pressuring the Pentagon in supporting the Afghan rebels with Stingers. The Soviet exposure did little to reduce American covert action in Afghanistan and instead increased American covert assistance to the Afghan rebels. As Operation Cyclone became transparent, Congress authorized the use of more overt and powerful action, namely the use of the Stingers. Thus, Soviet exposure led to American overt

escalation.

4 Comparative Statics

The previous section constructed a formal model of secret action and demonstrated that its implications capture empirical phenomenon. Our next task is to leverage the model to investigate how altering the environment changes the actors' equilibrium behaviors. In the process, we can examine whether existing theoretical claims withstand formal scrutiny for the types of cases covered under the model.

Indeed, we recover two unexpected findings. To begin, consider how adding value to secret action influences A's actions:

Proposition 4. *Within Proposition 3's parameter space, the probability that A takes secret action strictly decreases in the effectiveness of it (i.e., as V_S increases). Moreover, the overall probability of A ultimately implementing secret action strictly decreases in the effectiveness as well.*

This is bizarre. Increasing V_S makes secret action more powerful. Intuitively, A ought to pursue it more often. Many existing theories of secret action make this precise prediction. Significant increases in the ability to carryout secret action can encourage more use of secret action. For example, the creation of the CIA led to dramatic increases in the amount of covert action the US took during the early years of the Cold War (Rudgers, 2000; Johnson, 1989). Multiple former CIA operatives testified before the Church Committee regarding an agency which engaged in an over-reliance on covert action when limited, selective covert action is the intention of the law (Rudgers, 2000). The shift from military led covert action to an agency more capable of carrying out stronger covert action led to more covert action. Yet the opposite happens in equilibrium. Why?

What these papers do not take into account is the Robinson Crusoe fallacy (Tsebelis, 1989). Within Proposition 3's parameter space, the types of A sort themselves not by their own preferences but by B's. This is because A's strategy must induce B's indifference between exposing and not exposing. Otherwise, some types would certainly regret either taking secret action or not taking secret action.

The appendix proves the claim with an analysis of how V_S alters k^*. However, the unexpected result has a reasonable intuition. Within Proposition 3's parameter space, the types of A need to make B indifferent between exposing and not exposing. Increasing V_S makes B more inclined to expose A so as to force the higher cost types to quit and not enjoy the additional subversive power of secret action. Thus, to maintain B's indifference, exposing A must backfire more often conditional on having reached B's decision. Having fewer of the higher cost types—who would back down if exposed—engage in secret action accomplishes this. In effect, B's newfound desire to test the waters deters the marginal, least committed types and convinces them to accept the status quo. The probability that A takes secret action in the first place declines.

One may wonder whether the initial probability of secret action declines but the probability of secret action ultimately implemented increases. Despite the first half of Proposition 4, this is not immediately obvious. There are two ways A finishes firm with the secret action choice. To obtain either, A must enact the secret action in the first place. After that, Nature can not reveal the secret action, or Nature can reveal the secret action but B chooses not to reveal it.[12] The first process decreases in V_S as a consequence of the first half of Proposition 4. The second process decreases as well. As claimed a moment ago, larger values of V_S scare B and compel it to reveal the action more often. Thus, A becomes less likely to use secret action overall as it becomes more powerful.

Before moving on, we have a couple of technical points about Proposition 4. First, this comparative static increases the value of holding secret action while holding all other parameters constant. However, in the substantive motivation, we discussed how better secret action can also improve public action. In the appendix, we therefore replicate Proposition 4 under that assumption. Answering this question is more complicated than the main case and requires use of the implicit function theorem. Nevertheless, we derive a technical condition that guarantees the same implication. A sufficient condition for it is that the density of the type that indifferent between escalating and quitting is low.[13] This insures that the described intuition dominates the desire of more types to escalate.

Second, it is worth noting that local changes to V_S do not affect the probability of

[12] If B does reveal it, A either escalates to public action or backs down.

[13] That is, $f\left(\frac{V_P}{1-\alpha}\right)$ must be sufficiently small.

secret action for Proposition 1's parameter space. This is because all types pool on taking it. Large increases eventually shift the game to Proposition 3's case, however. This causes a decline in the probability of secret action similar to our central discussion, as the game goes from a situation where all types take it to a situation where some types do not. That said, further increases to V_S move the game to Proposition 2's parameters. There, types with k values less than $\frac{(1-p)V_S}{\alpha p}$ opt for secret action. This now causes the intuitive increase in the portion of types choosing to do so, as gambling on not getting caught looks more attractive. The aforementioned literature is on better theoretical ground for these parameter spaces.

Our second unexpected result concerns the probability B observes the action:

Proposition 5. *Within Proposition 3's parameter space, the probability that A takes secret action is unchanging in the probability Nature reveals the action. Moreover, the overall probability of A ultimately implementing secret action is also unchanging.*

One would suspect that increasing A's chance of exposure would deter it from taking secret action. Indeed, existing theories of covert action also make this prediction. Joseph and Poznansky (2018) show that exposure from information and communication technologies likely decreases traditional covert action. Likewise, security policymakers consider the risk of exposure when considering approving a covert action plan (Mark, 2015). Keefer (2001) also notes that policymaking institutions, like the Joint Staff, write reports on the risk of exposure for a potential covert action that is to be considered in the decision to carry out covert action. In the case of democratic leaders, Downes and Lilley (2010) highlights the that the decision to engage in covert intervention likely rests on the chance of success and a likelihood of a failure remaining secret. But we instead see that the frequency of covert action remains static. Why?

The answer here is more complicated than for Proposition 4, but the Robinson Crusoe fallacy is again at the heart of the problem. A's types choose their strategy intending to induce indifference from B. Note, however, that B's move occurs after Nature's decision to reveal the secret action. As a result, the probability of revelation does not determine whether B prefers exposing or not. In turn, A's types do not change their strategy in p. This explains why the probability of initial secret action remains fixed.[14]

[14]It also implies that if changing V_S necessarily increases the odds of revelation (perhaps because better secret action is inherently easier to observe), Proposition 4's relationship still holds.

Nevertheless, Proposition 5 goes further and says that the overall probability that A stays with the strategy is also identical. Given that mixed strategies normally intend to induce indifference and that the potential for exposure hurts A, one might expect that the end probability would be different. However, within Proposition 3's parameter space, B's strategy supplements the exposure risk. If the signal is weaker, B compensates by exposing more often; if the signal is stronger, B adds some slack by exposing less often. Inducing indifference requires that the changes perfectly offset one another. As such, the probability A stays with unexposed secret action does not change in p.

Like Proposition 4, these claims apply within Proposition 3's local region. Within Proposition 2's region, B exposes as a pure strategy. Thus, higher revelation rates deter more types from gambling on not getting caught. However, sufficient increases to p transition the game into Proposition 3, where the secret action rates stay flat thereafter.

5 Conclusion

This paper explored exposure of secret action. With no private information, such an interaction is straightforward. The opponent should expose if the state would give up and keep quiet if the state would escalate the situation. States that would escalate therefore choose to take secret action with impunity. Less resolved states may try their luck, hope that their secret actions remain secret, and then bow out if exposed. The least resolved states choose not to engage in secret action at all, deterred by the credible threat of exposure.

However, the exposer's dilemma becomes more complicated when faced with uncertainty over the state's resolve. Publicizing secret action becomes a gamble. If the exposer suspects that the state is very likely to double down, it withholds its knowledge. Less resolved states enjoy free rein here, conducting secret action they would otherwise not get away with. On the opposite end of the spectrum, if the exposer suspects that the state is very likely to give up, only the resistant types even try. Some bow out after exposure, but others escalate. The most interesting situation falls in between those cases, where some portion of unresolved types bluff strength by taking secret action. Without a clear response, the exposer sometimes publicizes the information and sometimes does not.

The bluffing behavior gives rise to a couple of unexpected comparative statics. Intuition would suggest that a state would become more likely to engage in secret action as secret action becomes more effective and others are less likely to observe it. However, such logic fails to properly account for the second-order incentives that come along with the opponent's decision whether to expose. The types that choose to take secret action do so in an effort to force the opponent into not having a clear response. Stronger secret action makes the opponent more inclined to try to end the action. Thus, more unresolved types must give up at the outset. Meanwhile, the opponent's exposure strategy equally counter-adjusts to any increase in the probability of revelation, meaning that secret action remains constant.

Zooming out, our model of uncertainty urges caution in interpreting observational evidence regarding the value of exposing and shaming perpetrators of secret action. Evidence of success from such exposure would suggest that policymakers should use the tactic more often. Lucas (1976) warns against making such recommendations based on historical data without an underlying theoretical explanation, and indeed our model highlights a problem. Strategic actors, to the best of their ability with the information they have available, choose expose questionable secret action when they believe that the results will be effective. We do not observe failures as often precisely because the actor endogenously chooses not to expose when it believes that the state will double down. In turn, more aggressive exposure without consideration for the strategic selection into secret action will backfire.

On a similar note, our model suggests a couple of paths forward for the secret action literature. What separates our exploration from Carson's (2016) is that the exposer tries to put the onus on the initiator to stop. Investigating uncertainty when a state wishes to keep a secret about another state's actions from its own population could also generate rewarding empirical implications. We also examined how the decision to expose affects selection into secret action in the first place, which is another worthy topic to explore further.

6 Appendix

We now prove the formal claims from the main text.

6.1 Proof of Proposition 1

Consider B's decision. If it exposes A, types with k values less than $\frac{V_P}{1-\alpha}$ escalate. Types with k values greater than that back down. The probability of the former case is the CDF of A's distribution evaluated at $\frac{V_P}{1-\alpha}$, or simply $F\left(\frac{V_P}{1-\alpha}\right)$. B then receives $-V_P$. The probability of the latter case is the complement. B then receives 0. Therefore, its expected utility for exposure equals:

$$-F\left(\frac{V_P}{1-\alpha}\right)V_P + \left(1 - F\left(\frac{V_P}{1-\alpha}\right)\right)0 = -F\left(\frac{V_P}{1-\alpha}\right)V_P$$

Meanwhile, B's utility for not exposing equals $-V_S$. As such, B does not expose if:

$$-V_S > -F\left(\frac{V_P}{1-\alpha}\right)V_P$$

$$F\left(\frac{V_P}{1-\alpha}\right) > \frac{V_S}{V_P}$$

This is the cutpoint in Proposition 1.

We only have A's strategies left to check for profitable deviations. By sticking to the strategy, A always secures V_S. This could be because Nature does not reveal the action or Nature does but B chooses not to expose it. Either way, V_S is the best possible payoff A can receive for the game, so cannot profitably deviate.

6.2 Proof of Proposition 2

Conditional on exposure, A's actions are a trivial application of backward induction. With that in mind, consider B's decision. If it exposes, types for which $k < \frac{V_P}{1-\alpha}$ escalate, and B receives $-V_P$. The remainder back down, and B receives 0. Given that only types with k values less than $\frac{(1-p)V_S}{\alpha p}$ take secret action in the first place, the posterior probability B is facing one of the types that will escalate is $\frac{F\left(\frac{V_P}{1-\alpha}\right)}{F\left(\frac{(1-p)V_S}{\alpha p}\right)}$. In turn, B prefers exposing to not if:

$$-\left(\frac{F\left(\frac{V_P}{1-\alpha}\right)}{F\left(\frac{(1-p)V_S}{\alpha p}\right)}\right)V_P + \left(1 - \frac{F\left(\frac{V_P}{1-\alpha}\right)}{F\left(\frac{(1-p)V_S}{\alpha p}\right)}\right)0 > -V_S$$

$$F\left(\frac{V_P}{1-\alpha}\right) < F\left(\frac{(1-p)V_S}{\alpha p}\right)\left(\frac{V_S}{V_P}\right)$$

This is the cutpoint in Proposition 2. Note that this parameter space implies that $\frac{V_P}{1-\alpha} < \frac{(1-p)V_S}{\alpha p}$. This is because $\frac{V_S}{V_P}$ is 0-to-1 constrained. If $\frac{(1-p)V_S}{\alpha p}$ exceeded $\frac{V_P}{1-\alpha}$, then

We only have A's strategies left to check for profitable deviations. But this is straightforward. Conditional on B always exposing, a type prefers to take secret action even if it were to back down after exposure if:

$$p(-\alpha k) + (1-p)V_S > 0$$
$$k < \frac{(1-p)V_S}{\alpha p}$$

Analogously, the remaining types would want to take no action.[15] This completes the proof.

6.3 Proof of Proposition 3

We begin by deriving B's indifference condition. If B does not expose, it earns $-V_S$. If it does expose, all types with k values less than $\frac{V_P}{1-\alpha}$ must escalate, and all types with k values greater than that must back down.[16] To derive a cutpoint strategy, suppose that all types with a value less than some k chose to engage in secret action in the first place. Then B is indifferent if:

$$-\left(\frac{F\left(\frac{V_P}{1-\alpha}\right)}{F(k)}\right)V_P + \left(1 - \frac{F\left(\frac{V_P}{1-\alpha}\right)}{F(k)}\right)0 = -V_S$$
$$\frac{F\left(\frac{V_P}{1-\alpha}\right)}{F(k)} = \frac{V_S}{V_P}$$

A unique solution exists. This is for the following reasons. To begin, note that the right hand side is between 0 and 1. Now consider the left hand side. Recall that the numerator is strictly positive, or $F\left(\frac{V_P}{1-\alpha}\right) > 0$. Thus, the right hand side goes to infinity as k approaches $\underline{k}$ from the right hand side. Meanwhile, as k goes to $\overline{k}$, the left hand

[15] Of course, types with $k < \frac{V_P}{1-\alpha}$ would want to escalate after exposure. But this only reinforces their desire to try secret action.

[16] What the $k = \frac{V_P}{1-\alpha}$ type chooses to do is immaterial because it has zero measure.

side goes to F $\left(\frac{V_P}{1-\alpha}\right)$. This is strictly less than the right hand side, otherwise we would be in Proposition 1's parameter range. Because the left hand side strictly decreases and is continuous in k, a unique value of k satisfies the line with equality. We call that value k^*. Note that $k^* > \frac{V_P}{1-\alpha}$, meaning that the k^* type prefers to back down following exposure. Moreover, $k*$ must also be greater than $\frac{(1-p)V_S}{\alpha p}$, otherwise we would fall into Proposition 2's parameter space.

Now consider what must be necessary for all types less than k^* to take secret action and all types less than k^* to not. Because the types' utilities are continuous in k, the type k^* must be indifferent between the two choices. The only way this can be true is if B mixes. Playing a pure strategy of not exposing means that A obtains V_S, which is strictly preferable to not taking secret action and earning 0. Meanwhile, playing a pure strategy of exposing means that the k^* type has a strict preference to maintain the status quo over taking secret action because $k^* > \frac{(1-p)V_S}{\alpha p}$. In turn, letting σ represent B's probability of exposing, the k^* type is indifferent if:

$$p(\sigma(-\alpha k^*) + (1-\sigma)V_S) + (1-p)V_S = 0$$
$$\sigma^* \equiv \frac{V_S}{p(V_S + \alpha k^*)}$$

These are the values in Proposition 3. The remaining strategies are straightforward applications of backward induction.

6.4 Proof of Proposition 4

We can prove the first half of Proposition 4 by examining Line 1. As V_S increases, the right hand side increases. Maintaining equality requires the right hand side to increase to compensate. The value k^* is the only degree of freedom. $F(k)$ strictly increases in k on the interval of the support of the distribution. Because this is in the denominator of the left hand side, the left hand side strictly decreases in k. As such, k^* must decrease when V_S increases.

For the second half of the proposition, note that the equilibrium probability of A implementing (and sticking with) covert action equals:

$$F(k^*)(1-p) + F(k^*)p(1-\sigma^*) = (1-p\sigma^*)F(k^*)$$

Recall that $\sigma^* = \frac{V_S}{p(V_S+\alpha k^*)}$. We can therefore further manipulate this probability to:

$$\left(\frac{\alpha k^*}{\alpha k^* + V_S}\right) F(k^*) \tag{2}$$

Because k^* decreases in V_S, $F(k^*)$ decreases as well. Thus, if $\frac{\alpha k^*}{\alpha k^*+V_S}$ also decreases in V_S, then the probability decreases overall. It will help to write k^* as a function of V_S for the corresponding derivative, giving us $\frac{\alpha k^*(V_S)}{\alpha k^*(V_S)+V_S}$. Showing that the derivative of this is negative gives us:

$$\frac{\alpha k^{*\prime}(V_S)(\alpha k^*(V_S) + V_S) - \alpha k^*(V_S)(\alpha k^{*\prime}(V_S) + 1)}{(\alpha k^*(V_S) + V_S)^2} < 0$$

$$k^*(V_S) > V_S k^{*\prime}(V_S)$$

The left hand side is obviously positive—the cutpoint k^* is always a positive value regardless of V_S. Meanwhile, the first half of this proof established that k^* decreases in V_S. This is the same thing as saying $k^{*'}(V_S) < 0$. But this means that the right hand side is negative. Overall, then, a positive value is greater than a negative value, thereby completing the proof.

6.5 Robustness Check for Proposition 4

We now derive the technical condition for Proposition 4 to hold if increasing the value of covert action also increases the value of public action. To do this, let $\epsilon > 0$ represent the change, such that the value for public action equals $V_P + \epsilon$ and the value of covert action equals $V_S + \epsilon$. Rather than take the derivative on V_S, we now want to know how changing ϵ alters the outcome to capture the effect of changing both covert and public action simultaneously.

Changing the notation for the extension and rearranging Line 1 yields:

$$\frac{F\left(\frac{V_P+\epsilon}{1-\alpha}\right)}{F(k)} - \frac{V_S + \epsilon}{V_P + \epsilon} = 0 \tag{3}$$

We want to know when increasing ϵ decreases the corresponding k that maintains the equality.[17] Such a condition implies that the probability of covert action decreases.

[17]From here, we require that a solution exists to Line 3 (given by assumption) and that the derivative of Line 3 with respect to k is non-zero (it is negative).

To do this, we can use the implicit function theorem on Line 3. The implicit function theorem tells us that the derivative of k^* with respect to ϵ is the negative of the derivative of Line 3 with respect to ϵ divided by the derivative of Line 3 with respect to k. That is, we require:

$$-\frac{\frac{\partial}{\partial \epsilon}\left(\frac{F\left(\frac{V_P+\epsilon}{1-\alpha}\right)}{F(k)}-\frac{V_S+\epsilon}{V_P+\epsilon}\right)}{\frac{\partial}{\partial k}\left(\frac{F\left(\frac{V_P+\epsilon}{1-\alpha}\right)}{F(k)}-\frac{V_S+\epsilon}{V_P+\epsilon}\right)}<0$$

Manipulating this yields:

$$-\frac{\frac{f\left(\frac{V_P+\epsilon}{1-\alpha}\right)}{F(k)(1-\alpha)}-\frac{V_P-V_S}{(V_P+\epsilon)^2}}{-\frac{f(k)F\left(\frac{V_P+\epsilon}{1-\alpha}\right)}{F(k)^2}}<0$$

$$f\left(\frac{V_P+\epsilon}{1-\alpha}\right)<\frac{F(k)(1-\alpha)(V_P-V_S)}{(V_P+\epsilon)^2}$$

The right side is strictly positive. Thus, a sufficiently small density for the type indifferent between escalating and quitting generates the result.

6.6 Proof of Proposition 5

The proof here is straightforward. The probability of A taking covert action is $F(k^*)$, and the probability of A ultimately implementing it is Line 2. Because p does not appear explicitly in either of these figures, for either probability to change in p, it must be that k^* is implicitly a function of p. However, Line 1 shows that this is not the case, thereby completing the proof.

References

Al Arabiya. 2012. "U.S. officials confirm captured Iranians in Syria are active military: report." *Al Arabiya* . 1

Al Jazeera. 2012. "Iran 'sending arms to Syria despite ban'." *Al Jazeera* . 1

Beeley, Harold. 1962. "Harold Beeley, ambassador, Cairo to Foreign Office,.". 11

Brzezinski, Zbigniew. 1979. "Reflection on Soviet Intervnetion in Afghanistan.". GWU. 21

Butler, Rab. 1964. "Rab Butler to Prime Minister.". 11

Carson, Austin. 2016. "Facing off and saving face: Covert intervention and escalation management in the Korean War." *International Organization* 70(1):103–131. 1, 3, 10, 22, 27

CIA. 1985. "Scope: Hostage Rescue- Middle East.". GWU. 15

Cogan, Charles G. 1993. "Partners in Time: The CIA and Afghanistan since 1979." *World Policy Journal* 10(2):73–82. 22

Cormac, Rory and Richard J. Aldrich. 2018. "Grey is the new black: covert action and implausible deniability." *International Affairs* 94(3):477–494. 5, 12

CoSC. 1964. "Top Secret: Chief of Staff Committee: Aden and the South Arabian Federation.". 11

Dagher, Sam and S Fitch. 2015. "Iran expands role in Syria in conjunction with Russias airstrikes." *Wall Street Journal* . 1

Dai, Xinyuan. 2005. "Why comply? The domestic constituency mechanism." *International Organization* 59(2):363–398. 5

DOS. 1986. "Brunei Project.". GWU. 15

Downes, Alexander B. and Mary Lauren Lilley. 2010. "Overt Peace, Covert War?: Covert Intervention and the Democratic Peace." *Security Studies* 19(2):266–306. 25

Edwards, Julia. 2015. "Obama acknowledges damage from NSA eavesdropping on Merkel." *Reuters* . 4

Fathollah-Nejad, Ali. 2018. "Iranians respond to the regime: 'Leave Syria alone!'." *AlJazeera* . 4

Fey, Mark and Kristopher W Ramsay. 2011. "Uncertainty and Incentives in Crisis Bargaining: Game-Free Analysis of International Conflict." *American Journal of Political Science* 55(1):149–169. 7

Filkins, Dexter. 2013. "The Shadow Commander: Qassem Suleimani is the Iranian operative who has been reshaping the Middle East. Now hes directing Assads war in Syria." *The New Yorker* . 1

Finnemore, Martha and Kathryn Sikkink. 1998. "International norm dynamics and political change." *International organization* 52(4):887–917. 5

Finney, John. 1973. "More Secret Air Attacks In Cambodia Disclosed." *The New York Times* p. 1. 5

Fisher, Dan. 1980. "Kremlin Steps Up Anti-U.S. Campaign, Charges CIA Is Training Afghan Rebels." *Los Angeles Times* p. B8. 22

Fisk, Robert. 2013. "Iran to send 4,000 troops to aid President Assad forces in Syria." *The Independent* 16. 1

Gibbs, David N. 2000. "Afghanistan: The Soviet invasion in retrospect." *International Politics* 37(2):233–245. 21

Greenwood, Max. 2018. "US sanctions five Iranians over missile support to Houthis." *The Hill* . 4

GRU Memorandum. 1988. "About US Aid to the Counterrevolutionaries (Excerpt).". 22

Hansard. 1964. "Contains Hansard parliamentary answers.". 12

Harsanyi, John C. 1973. "Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points." *International journal of game theory* 2(1):1–23. 21

Inouye, Daniel and Lee Hamilton. 1987. Report of the Congressional Committees Investigating the Iran-Contra Affair. Congressional 100-216 US Congress Washington: . 15

Isenberg, David. 1989*a*. "Cato Institute Policy Analysis No. 118: The Pitfalls of US Covert Operations.". 21

Isenberg, David. 1989*b*. "The Pittfalls of US Covert Operations.". **URL:** *https://object.cato.org/pubs/pas/PA118.HTM* 5

JCS. 1962. "Justification for US Militarry Intervention in Cuba.". GWU. 13, 14

Johnson, Loch K. 1989. "Covert Action and Accountability: Decision-Making for America's Secret Foreign Policy." *International Studies Quarterly* 33(1):81–109. 23

Joseph, Michael F and Michael Poznansky. 2018. "Media technology, covert action, and the politics of exposure." *Journal of Peace Research* 55(3):320–335. 1, 4, 25

Keefer, Edward C. 2001. "Foreign Relations of the United States, 1964–1968." *Volume XXVI. Indonesia; Malaysia-Singapore; Philippines* . 25

Kuperman, Alan J. 1999. "The Stinger Missile and U.S. Intervention in Afghanistan." *Political Science Quarterly* 114(2):219–263. 22

Lebovic, James H and Erik Voeten. 2009. "The cost of shame: International organizations and foreign aid in the punishing of human rights violators." *Journal of Peace Research* 46(1):79–97. 5

Lewis, Anthony. 1976. "Menu for Disaster." *The New York Times* p. 21. 5

Lucas, Jr., Robert E. 1976. Econometric policy evaluation: A critique. In *The Phillips curve and labor markets*, ed. Karl Brunner and Allan H Meltzer. North Holland: Elsevier. 27

Macmillan, Harold. 1962. "Prime Minister personal minute.". 11

Mark, Lowenthal. 2015. "Intelligence: From Secrets to Policy.". 25

McFarlane, Robert. 1985. "Recommend Telephone Call.". GWU. 15

McNamara, Robert. 2017. "The Nasser factor: Anglo-Egyptian relations and Yemen/Aden crisis 196265." *Middle Eastern Studies* 53(1):51–68. 11

McVeigh, Karen. 2017. "'Trumps secret Yemen war': UK role in US counter-terrorism causes unease." *The Guardian* . 4

Newton, Robert. 1992. "The Capture ofthe USS Pueblo and Its Effect on SIGINT Operations.". 5

North, Oliver. 1986. "Release of American Hostages in Beirut.". 15

Orkaby, Asher. 2017. "The North Yemen civil war and the failure of the Federation of South Arabia." *Middle Eastern Studies* 53(1):69–83. 11, 12

OSD. 1962. "Memorandum for the Record.". 13, 14

Pach, Chester. 2006. "The Reagan Doctrine: Principle, Pragmatism, and Policy." *Presidential Studies Quarterly* 36(1):75–88.
URL: *http://www.jstor.org/stable/27552748* 21

Paine, Jack and Scott A Tyson. 2019. "Uses and abuses of formal models in political science." Manuscript, University of Rochester. 9

Pear, Robert. 1988. "Arming Afghan Guerrillas: A Huge Effort Led by U.S." *The New York Times* . 22

Rudgers, David F. 2000. "The Origins of Covert Action." *Journal of Contemporary History* 35(2):249–262. 23

Sagar, Rahul. 2001. *Whistleblowers: Broken lives and organizational power*. Ithaca: Cornell University Press. 6

Sagar, Rahul. 2013. *Secrets and leaks: The dilemma of state secrecy*. Princeton: Princeton University Press. 6

Spaniel, William and Michael Poznansky. 2018. "Credible Commitment in Covert Affairs." *American Journal of Political Science* 62(3):668–681. 3, 21

Tehran Times. 2017. "Iran fires missiles at ISIS in Syria for Tehran attacks." *Tehran Times* . 1

Terman, Rochelle. 2019. "Rewarding resistance: Theorizing defiance to international shaming." Manuscript, University of Chicago. 6

Thompson, Nick. 2017. "Ukraine: Everything you need to know about how we got here." *CNN. com. Last modified February* 3. 4

Tisdall, S. 2012. "Iran helping Syrian regime crack down on protesters, say diplomats [Electronic resourse]." *The Guardian* . 1

Tsebelis, George. 1989. "The abuse of probability in political analysis: The Robinson Crusoe fallacy." *American Political Science Review* 83(1):77–91. 23

Ustinov, Dimitri F. 1980. "Report by Soviet Defense Minister Ustinov to CPSU CC on 'Foreign Interference' in Afghanistan.". 22

Weinberger, Caspar. 1985. "Weinberger Diary.". GWU. 15, 16

Weinberger, Caspar. 1986. "Meeting ... with the President ... in the Oval Office.". GWU. 15, 16

Wolford, Scott. 2007. "The turnover trap: New leaders, reputation, and international conflict." *American Journal of Political Science* 51(4):772–788. 6