Secret Innovation[*]

Michael F. Joseph[†]   Michael Poznansky[‡]

May 20, 2022

Abstract

Conventional wisdom holds that open, collaborative, and transparent organizations are innovative. But some of the most radical innovations—satellites, lithium-ion batteries, the internet—were created by small, secretive teams in national security agencies. Are these organizations more innovative because of their secrecy or in spite of it? We introduce political preferences and secrecy into a principal-agent model of organizational innovation. We argue that secrecy facilitates innovation for high-risk, high-reward projects through a cost-passing mechanism. In open institutions, managers will not approve pilot research into high-risk, high-reward ideas for fear of incurring political costs. Researchers exploit secrecy to conduct pilot research at a higher personal cost to generate evidence that their project is viable and win their manager's approval. Contrary to standard principal-agent findings, we show that researchers may exploit secrecy even if their preferences are perfectly aligned with their manager's; and that managers do not monitor researchers even if monitoring is costless and perfect. We illustrate our theory on two cases from the early Cold War: the successful development of the U-2 spy plane and the CIA's failed foray to master mind control (MK-ULTRA). We contribute to the political application of principal agent theory, and studies of political innovation, conflict and innovation, emerging technologies, abuse and oversight, and great power competition.

1   Introduction

Nations that want economic prosperity, sound public policy, and enhanced national security must innovate (Black and Lynch, 2004; Freeman, 2015; Horowitz, 2010*a*; Taylor, 2016). To wit, scholars across many disciplines study why some organizations are more innovative than others. There is broad agreement that openness, defined loosely as organizations which encourage employees to share their ideas with their colleagues, spurs innovation (Bresnahan, Brynjolfsson and Hitt, 2002; Laursen, 2003; Tushman and Anderson, 1986; West and Anderson, 1996).[1] Open organizations foster competition and collaboration between otherwise siloed divisions (Aghion, Bloom, Blundell, Griffith and Howitt, 2005; Macdonald, 2015), diffuse ideas (Boushey, 2016), and encourage a free-flow of information (Zoghi, Mohr and Meyer, 2010).

One set of institutions consistently buck this trend: secretive intelligence and national security organizations. Agencies like the Central Intelligence Agency (CIA), the National Reconnaissance Office (NRO), and MI5 are notoriously siloed. Usually only a few scientists directly involved in a research project know the details.[2] And yet, secretive national security organizations have produced radical innovations including the satellite,[3] artificial intelligence and speech recognition programs,[4] autonomous robots,[5] lithium-ion batteries,[6] nuclear weapons, the internet, component-part factories,[7] and GPS and other systems that birthed Google Earth.[8] They also revolutionized microchips, wind turbines, and other renewable energy technologies.[9] Failed projects speak to the scope of their work as well. During World War II, the Office of Strategic Services (OSS) devised a plan to cover foxes in glow-in-the-dark paint to scare Japanese soldiers. As proof of concept, they dropped glowing foxes into Central Park, terrorizing New Yorkers (Houghton, 2019). The CIA also explored psychic spying for 40 years to determine if individuals claiming paranormal abilities could reveal Soviet capabilities (Richelson, 2002, 176-187).[10]

---

[1]For exceptions, see Rotemberg and Saloner (1994).

[2]Senior managers sometimes know some details of particular projects, but not always.

[3]https://www.nro.gov/Portals/65/documents/history/csnr/corona/The%20CORONA%20Story.pdf?ver=BgSn5nPYz45EZ9O$_Z$F5

[4]https://dc.mit.edu/sites/default/files/pdf/2014-federally-supported-innovations.pdf, pp. 13-14

[5]https://www.nist.gov/publications/learning-hierarchical-control-system-4drcs-darpa-lagr-program

[6]https://www.cia.gov/about-cia/cia-museum/experience-the-collection/text-version/stories/cias-impact-on-technology.html.

[7]This significantly advanced the industrial revolution (Winchester, 2019, 86-90).

[8]https://www.youtube.com/watch?v=XTCoOKZoDIElist=PL979C3B52F202C43Findex=4.

[9]https://stacker.com/stories/5483/50-inventions-you-might-not-know-were-funded-us-government

[10]The Federation of American Scientists reviewed classified 'remote viewing' programs. They found evidence that some operators likely held psychic abilities. However, these psychic abilities were not sufficiently accurate to be

Are these organizations more innovative because of secrecy, or in spite of it? To answer this question, we study a principal-agent model of organizational innovation (e.g Lai, Riezman and Wang, 2009; Kopel and Riegler, 2006). We adapt the payoffs to reflect the agents' sensitivities to political costs and benefits (Joseph, Poznansky and Spaniel, 2022). We then contrast mechanisms for innovation in open and secret public-sector institutions (Cain, 2014).[11]

Secrecy allows different agents to distribute the political costs of authorizing each phase of controversial, high-risk research programs. In open institutions, lower-level researchers cannot pursue pilot programs to determine if a project is viable without their manager learning about it. When their manager learns of a novel but controversial idea, she will not even approve pilot research to determine if the project is viable because she does not want to be responsible for the decision. Secrecy gives an enterprising researcher cover to collect evidence (at a larger personal cost) that the novel idea is viable. When a pilot study shows promise, the researcher can take it to their manager for approval.

Including secrecy and political preferences generates two surprising results for principal-agent theory (Downs and Rocke, 1994; Miller, 2005; Hawkins, Lake, Nielson and Tierney, 2006; Di Lonardo, Sun and Tyson, 2020). First, the researcher turns to secrecy even if her preferences are perfectly aligned with the manager's preferences. Second, the manager does not monitor the researcher even if monitoring is costless and perfect *and* the manager knows that the researcher only exploits secrecy to do something the manager would not allow her to do. These results follow from a don't-ask-don't-tell dynamic made possible because secrecy allows actors to distribute costs. Distributing costs alleviates preference asymmetry. The manager knows if she monitors the researcher, she will discover something unsavory and shut research down. However, if she remains ignorant, she can incur a small share of the costs associated with highly controversial pilot research, and still benefit when the pilot research shows promise.

We also use the model to explore the qualities of innovations unique to secretive, national security institutions. We find that all political organizations pursue ideas that, in expectation, serve the national interest and involve non-controversial research practices. However, only secret organizations can pursue initial concepts that involve large risks and rewards. This includes promising

---

useful. See https://irp.fas.org/program/collect/air1995.pdf.

[11]We emphasize *internal* secrecy where small teams can conceal their practices from their colleagues, managers, etc. This is distinct from *external* secrecy wherein states are concealing information from other countries.

ideas that require controversial pilot research. It also includes obscure ideas that could generate either enormous gains or losses once deployed in the field. Before research is carried out, these ideas are too controversial for open, public-sector organizations to pursue. With hindsight, they represent both the path-breaking innovations the intelligence community is known for, and some of its shameful failures.

We illustrate our theory using two cases: innovation to facilitate reconnaissance missions over the Soviet Union (resulting in the U-2 spy plane); and attempts to master mind control (the highly controversial MK-ULTRA program). While these cases are known to intelligence historians,[12] their implications have not been explored for political science scholarship interested in related cases (Coe and Vaynman, 2019; Vaynman, 2022; Carnegie and Carson, 2018).

At the broadest levels, we contributes to studies of organizational innovation and principal-agent dynamics (Laursen, 2003; Hirsch, 2016; Foarta and Sugaya, 2021; Hawkins et al., 2006). We contribute to studies of political innovation by detailing and then modeling preferences and costs that are unique to political agents (Boushey, 2016; Breznitz and Ornston, 2013; Frant, Berry and Berry, 1991; Gray, 1973; Mintrom, 1997; Taylor, 2016). We also contribute to four debates in international relations and comparative conflict. First, we broaden arguments about innovation in autocratic regimes and terrorist organizations by showing how secrecy allows democratic governments to pursue politically sensitive ideas (Dolnik, 2007; Dragu and Lupu, 2021; Gohdes, 2020; Moghadam, 2013; Perkoski, 2019; Horowitz, 2010b). Second, some wonder why the military often finds innovation difficult even though national security is important (Farrell and Terriff, 2002; Posen, 1984; Rosen, 1988).[13] We expand the scope to other national security institutions and show that they are prolific innovators because they are internally secret. Third, we advance a national-security theory of bottom-up innovation (Griffin, 2017; Grissom, 2006; Foley, 2012; Jungdahl and Macdonald, 2015; Macdonald, 2015). Fourth, we provide one mechanism for balancing positive and negative ethical and strategic implications of emerging national security technologies (Horowitz, 2016; Zhang, Anderljung, Kahn, Dreksler, Horowitz and Dafoe, 2021; Sechser, Narang and Talmadge, 2019). A detailed review of the security literature and additional implications is provided in Appendix D. We discuss policy implications in the conclusion.

---

[12]e.g. Andrew (1995); Bateman (2020); Hopkins (1996); Doel and Needell (1997); McCarthy (2013); Richelson (2002); Wallace, Melton and Schlesinger (2009).

[13]See also Griffin (2017, 215).

2 Concepts

Our theory is closest to principal-agent models that examine rationalist, organizational innovation (e.g Lai et al., 2009; Kopel and Riegler, 2006). We adapt this model to account for political preferences and national security dynamics. Others introduce political preferences into principal-agent problems (see Miller, 2005, for review). But they typically focus on policy-making or electoral accountability (Downs and Rocke, 1994, is closest to this study); and not innovation. Some examine principal-agent problems that are unique to international relations. But they emphasize interactions between two states (Hawkins et al., 2006), or two foreign militaries (Biddle, Macdonald and Baker, 2018). We focus on a handful of employees working within government agencies. We detail the differences in Appendix E. Here we examine three important features of our principal-agent problem: organizational innovation; political costs and benefits; and secrecy.

2.1 Innovation and its effects

Innovation is the process of taking a novel idea and converting into a working device or policy (Kollars, 2017, 126).[14] Innovation occurs only after (1) a novel idea, (2) pilot testing to improve and validate that insight, and (3) the decision to develop a product and deploy it in the field (King, 1990). The last step is critical. It is not enough to conceive an idea. Innovation requires that the idea is developed into a working product (West and Anderson, 1996).

Government agencies innovate to achieve their policy goals (Taylor, 2016). National security goals include monitoring rivals or preventing terrorist attacks. We define an innovation's effects as whether the final product moves the nation towards or away from its goals. Many innovations have positive effects (i.e. move the organization towards its goals). Others have no effect. Others still have negative effects because of unintended consequences (Sechser et al., 2019). This could include conflict escalation, degrading defenses, or facilitating local rebellion (Horowitz, 2020; Kuo, 2020). While researchers hold expectations, they are uncertain about the true effect.

---

[14]Taylor (2016, 29) defines innovation as "as the discovery, introduction, and/or development of new technology, or the adaptation of established technology to a new use or to a new physical or social environment."

2.2 Political costs from developing new ideas

We distinguish between the effects of innovation that follow from deploying a product (described above) and the costs associated with moving an idea through the development phases (King, 1990). Some development costs stem from the financial burden of research trials and prototype construction. But public-sector institutions are especially sensitive to political costs.[15] Political costs stem from professional punishments, moral costs, or garnishes imposed upon agents following a policy choice that is morally questionable, embarrassing, or perceived as a waste of government resources (Caillier, 2017; Colaresi, 2014).[16] These political costs can manifest at different stages of innovation. During pilot research, political costs can be activated from wasteful spending or human subjects research without consent (Sagar, 2013). During the deployment phase, political costs can be activated from labor abuses during production, or political fall-out—such as escalation risks with foreign rivals or damage to international reputation—from revealing a controversial project.

Of course, not all research activates political costs.[17] But in many cases, national security employees do face costs for pursuing ideas. National security employees often face large costs for innovation because of an organizational culture that perceives radical ideas as reckless, steep punishments for perceived abuse of public trust, and bureaucratic inertia (Grissom, 2006; Lee, 2019; Price, 2014).

2.3 Internal secrecy determines who bears costs

Secrecy allows individuals to make choices without first seeking approval from superiors and without those superiors and the public learning about their actions for a long period of time (Cain, 2014). We focus on secrecy over pilot research: a researcher's capacity to develop a prototype, run laboratory tests, simulations, or other research programs without a manager or compliance officer knowing about it.

Secrecy is similar to the delegation of authority (Laursen, 2003); but there are important dif-

---

[15]We accept that political and financial costs exist in private and public sectors. However, profits are the main focus of private sector innovation (see Freeman, 2015; King, 1990).

[16]For instance, public sector employees face unique punishments such as jail time associated with mishandling public money. National security employees are held to the highest professional standards through their security clearances. Evidence shows that public sector employees have strong public service motivations than private sector employees (Caillier, 2017).

[17]Our model accounts for this because we allow costs to be 0.

ferences. Studies of private-sector innovation show that innovative firms do delegate spending and research choices to subordinates (Bresnahan et al., 2002). But delegation works because teams collaborate and compete with other teams in the organization (Jones, Kalmi and Kauhanen, 2006; West and Anderson, 1996; Aghion et al., 2005), or take other actions that can only happen if research is open (Zoghi et al., 2010; Laursen, 2003; Bresnahan et al., 2002).

Unlike delegation, secrecy ensures that researchers can spend years on controversial projects without anyone learning the details of what they are doing (Cain, 2014).[18] For example, the United States' federal budget is designated at the program level, and monitoring and evaluation is mandatory for *open* agencies. Spending choices are subject to external evaluation so the government can verify public funds are well spent. In contrast, secretive intelligence agencies and parts of the military have access to an "Unvouchered Contingency Reserve Fund" that allows them to spend money without explaining what it is for.[19]

Also unlike delegation, secrecy allows managers to avoid political costs through ignorance. When a scandal erupts in an *open* government organization, a manager cannot easily say they did not know what their staff was doing because the public expects them to monitor their employees. But national security employees are expected to maintain secrecy to guard against leaks, and counter-intelligence threats. This helps excuse managers who do not intrusively monitor their staff to learn about questionable choices. For example, during the Iran-Contra Affair Reagan avoided significant costs by claiming that subordinates engineered the scheme to use proceeds from arms sales to Iran to fund the Contras in Nicaragua without his knowledge.[20]

Of course, scientists that authorize research in secret organizations still keep detailed project records that managers can request. Managers also have a top-level understanding of a project's objectives. But managers have no incentive to inquire about project details. Often the devil is in these details. In the 1960s, a secret military research program tested the effects of various chemicals agents on U.S. soldiers. The project received approval under the assumption that subjects provided consent.[21] An Inspector General report later found, however, that "volunteers were not

---

[18]Cain argues that secrecy allows all elites to avoid costs in finding compromise in a policy bargaining context. We examine how secrecy allows agents to distribute costs of making choices (not raising ideas to reach a compromise) in a principal-agent context.

[19]https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/agency-and-the-hill/09-The%20Agency%20and%20the%20Hill_Part2-Chapter6.pdf

[20]In contrast, Secretary of Defense Casper Weinberger was indicted on criminal charges.

[21]https://nsarchive2.gwu.edu/radiation/dir/mstreet/commeet/meet4/brief4.gfr/tab_l/br4l1a.txt.

fully informed, as required, prior to their participation."[22] The project went on for years before managers learned of this detail and the project was shut down.[23]

In practice, agents can exploit secrecy at different levels of a secret organization. To keep things simple, we detail a two-level institution that involves one decision-maker and one researcher. However, in many historical examples we see variation between who knows the devilish details and who does not. At one extreme, a handful of scientists know the controversial details of a program, but even their immediate superiors are unaware of the controversial research activities. At the other extreme, the president is fully aware of the devilish details, but Congress is not. In the middle, directors of intelligence agencies know exactly what their subordinates are doing but do not inform the president.[24] If we add layers of management to the institution, our basic predictions still bear out so long as there is secrecy at some level of the organizational hierarchy. There must be at least one partition between insiders who can pursue research and development without explaining their practices outside of the group and who share the costs of authorization if things go wrong, and outsiders who can save some costs by remaining ignorant about what her subordinates are up to but cannot stop programs for a long time.

3   Model

Our analysis plan is as follows. First, we set-up a basic institution. Second, we formally define secret innovation, and contrast the process of innovation in secret and open organizations to explain the core mechanism that drives secret innovation. Third, we use comparative statics to explore the innovations uniquely pursued in secret organizations. Fourth, we introduce two distinct information, agency, and monitoring problems into the model to flesh-out the mechanism and connect the model to the principal-agent literature.

3.1   Setup

We study an institution that employs two agents: a researcher (R, she) and a manager (D, for decider, he). Figure 1 visualizes the game-tree and payoffs. The dashed box is the sub-game in

---

[22]https://bioethicsarchive.georgetown.edu/achre/final/chap3_4.html.

[23]See p. 411 of https://www.intelligence.senate.gov/sites/default/files/94755_I.pdf.

[24]In other examples there are inter-agency teams. But the teams are small and secret. Our theory covers any project team that can maintain secrecy, whether all members work for the same agency or not.

which R exploits secrecy. In it, she can conduct pilot research without her manager knowing about it.[25] Below we contrast secret and open institutions. Open institutions remove the secret sub-game but are otherwise identical.

We model the true effect of unleashing a new innovation on the world as $\pi \in \mathcal{R}$. When π is positive (negative), it means that the innovation ultimately moves the institution closer (further) from achieving its goals. Of course, agents cannot anticipate all the consequences of unleashing new devices on the world ex-ante. Thus, D's choice to innovate is based on an expectation of the consequences. Define $p(\pi) \to \mathcal{R}$ as a density function that determines the effect of introducing the innovation into the world. We assume that both player's know the density function $p()$, but not the true realization of π.

Along the way to innovation, agents can authorize pilot research, which has two effects. First, pilot research improves the value of innovation by $\theta \geq 0$. Second, pilot research helps discover the true effect if innovation happens. We model this as a normally distributed signal $m \sim \mathcal{N}(\pi, \sigma)$ tied to the true consequences of innovation (π).[26]

Agents pay political costs for participating in a controversial research process. We assume that players pay one cost—$k_i, i \in \{R, D\}$— if the institution engages in pilot research.[27] They pay a second cost—c_i—if the project is deployed into the field. We assume that actors incur costs based on how responsible they are during the decision-making process. The total amount of cost to be apportioned in $1 + x$. We distribute 1 unit of cost to the agent that chooses to take costly action (conduct research, authorize innovation), and a smaller $x \in (0, 1)$ portion to the other agent who works at the same institution but who did but did not directly take a costly action.

3.2 Analysis: Secret innovation and the cost-passing mechanism

Our solution concept is sub-game perfect equilibria (SPE). We define secrete innovation as follows.

Definition Suppose parameter values in the open institution where innovation does not occur with probability on the path in any SPE. Then **secret research facilitates innovation** if innovation

---

[25]We study more complex affects of secrecy in extensions.

[26]Note π is drawn from an arbitrary distribution. We model the signal from a normal distribution to avoid corners if $p(\pi)$ is supported on a limited range.

[27]A cost of $k = 0$ implies that pilot research isn't controversial.

Figure 1: Game Tree
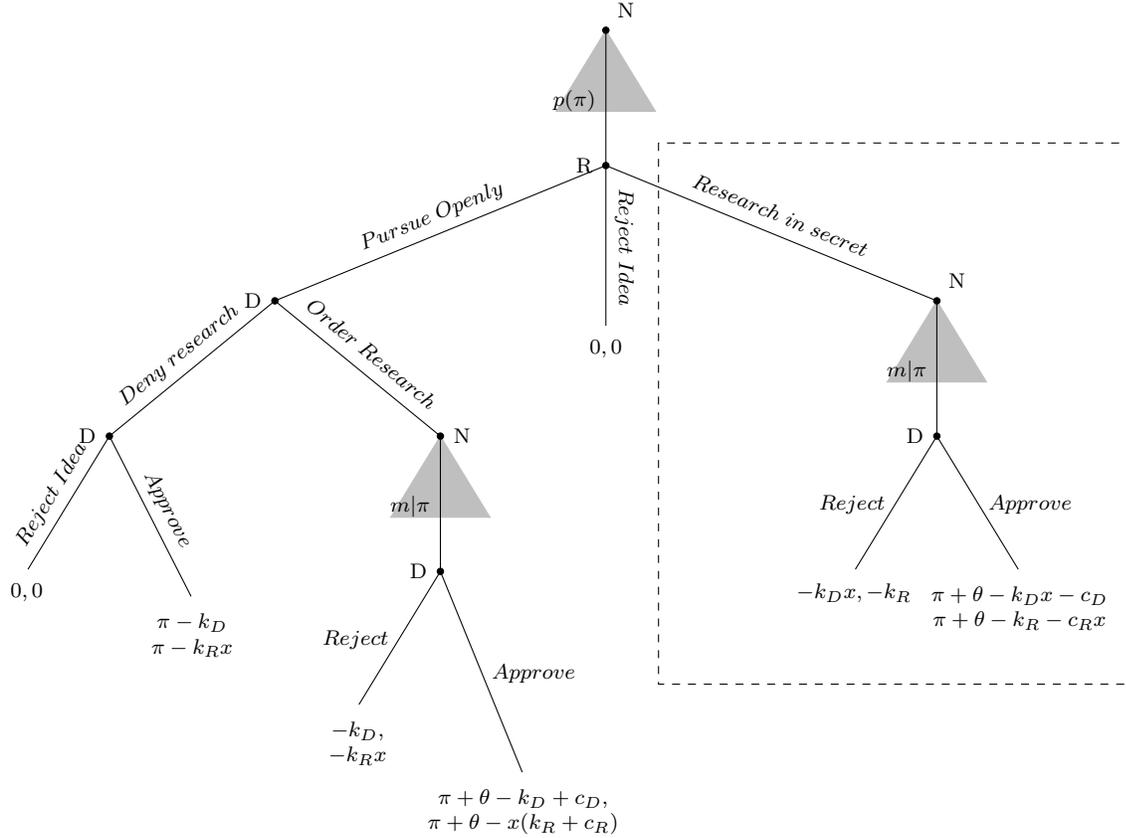
The dashed rectangle represents the secret option. The counter-factual open organization does not include this sub-game.
Shaded triangles represent random variables. Nature does not reveal π to either player. Nature reveals m to both players.

Parameter	Interpretation
$\pi \in \mathcal{R}$	Affect from unleashing an innovation on the world
$p(\pi)$	Agents' initial expectation about the innovation's affect.
$k_i \geq 0$	Political costs agent i incurs when innovation is approved
$c_i \geq 0$	Political costs agent i incurs when pilot research is approved
$x \in [0,1)$	What the agent pays for participating in innovation phase that she did not authorize
$m()|\pi$	What agents learn from pilot research
$\theta \geq 0$	How much pilot research improves an idea

occurs with positive probability in any equilibrium in the secret institution with the same parameter
values.

This definition highlights the counterfactual nature of our claim. Open institutions can innovate.
But there are some ideas that only secret institutions will pursue.

Our first task is to identify the ideas open institutions will not pursue. Define $e_0|p$ as the agents'
prior expected value for π. Define, $e_1|p,m$ as the agents' posterior expected value for π if research
happens and agents observe a message m. Finally, define $\lambda_0 = pr(e_1 > k_D - \theta)|p$ as D's pre-research

belief that if research is conducted, he will observe a signal m that will lead to a posterior belief $e_1 > k - \theta$. This expectation is essential. The reason is that once research is complete, D is willing to authorize a project if and only if: $e_1 + \theta - c - k > -c \equiv e_1 > k - \theta$. When λ_0 is low it means that D believes that even if he allows for controversial research he is still unlikely to approve the project because the research will show the project will not work. These expectations will help us bound the conditions for innovation in the open institution.

Lemma 3.1 *Neither research nor innovation can happen in the open institution if*

$$\lambda_0 < \frac{c_D}{e_0 + \theta - k_D} \tag{1}$$

In every SPE player utilities are $U^D = U^R = 0$.

See Appendix A.1. When condition 1 is satisfied, D never innovates or even conducts research to determine if the project is viable. Two factors drive D to reject a request for pilot research. First, research involves political costs (c).[28] Second, at the point where D is asked to authorize controversial research, his expectation about that research is inextricably connected to his prior belief. When preexisting scientific research suggests the project is not promising, D expects future research to, on average, confirm that expectation.

We now turn to the secret institution. Since we are interested in the cases where secrecy facilitates innovation, we focus on the conditions where innovation cannot happen in the open institution.

Proposition 3.2 *Secrecy facilitates innovation if condition 1 and:*

$$\frac{c_R}{e_0 + \theta - k_R x} < \lambda_0 \tag{2}$$

are satisfied. If they are, then in every SPE, R exploits secrecy to conduct pilot research, D authorizes the project if and only if that research provides evidence the program will work. Off the path, if R attempts to pursue open research, D denies R's research and innovation does not happen.

See Appendix A.2. The result describes a condition where the researcher is willing to exploit secrecy to conduct research (condition 2 is satisfied), but her manager was unwilling to approve

---

[28] Trivially, if research is costless or beneficial you always see open innovation.

open research (condition 1 is satisfied). If research provides evidence that the project is viable (m suggests π is higher than originally thought) then the manager will approve the project, leading to an innovation.

Notice that we can achieve secret research even if the manager and the researcher's cost functions are identical: $k_R = k_D, c_R = c_D$. This is surprising given what we know about principal-agent problems. In standard accounts, researchers only exploit secrecy when their preferences diverge from the manager. Why is a researcher with the same incentives as the manager willing to conduct research when her manager is not? The answer comes down to cost passing. Secrecy gives the researcher discretion to conduct pilot research to try to convince the manager, who is unwilling to pay the research costs, to approve if it shows promise. If R's secret pilot research shows promise (m is large) she can take the results to her manager for approval. Thus, the researcher is willing to take on the up-front cost and risk of research because she can convince her manager to bear the brunt of the deployment cost.

3.2.1 Predictions about ideas: Secrecy drives innovation when initial ideas are high-risk, high-reward

What are the kinds of initial ideas researchers need secrecy to pursue? Using a comparative static analysis, we expose two ideal-type pathways to secret innovation that are made possible because the manager and the researcher weigh certain trade-offs differently. We provide technical support for these two pathways in Appendix A.3. We visualize the results in Table 1. These pathways can interact. However, the basic trade-offs that we identify are always present. Thus, it is valuable to consider them as distinct.

The first pathway appreciates the agent's initial expectations about whether an idea will provide a benefit ($p(\pi)$). In real life, a researcher uses publicly available research on related problems to make predictions about what will happen if her idea is developed. Column 1 of Table 1 plots the initial expected consequences of four different concepts institutions could pursue. Row 1 is the baseline. The remaining three panels represent different ways initial beliefs can vary.

First, they vary in their on average, expected effects (e_0). As e_0 increases (row 2), it means that the institution's initial expectation is that the idea is increasingly likely to yield a net benefit if it is developed and deployed into the field.

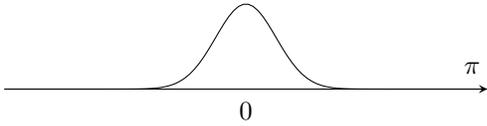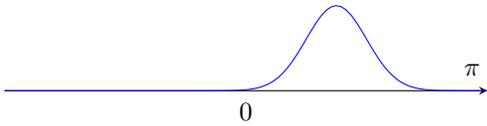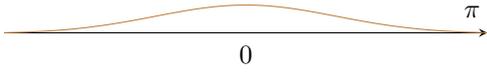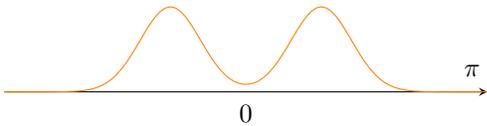Table 1: The innovation pathways for different initial ideas

Initial predictions of effects: $p(\pi)$	Substantive description	c_D low	c_D high
Baseline: $e_0 = 0.1, \sigma_0 = 1$			
[distribution plot, peak near 0, axis labeled π, origin 0]	Project team agrees, with little variance, that innovation will not impact state goals.	Scrap idea	Scrap idea
More Optimistic, same confidence: $e_0 = 3, \sigma_0 = 1$			
[distribution plot, peak right of 0, axis labeled π, origin 0]	Project team agrees, with little variance, that innovation will positively impact state goals.	Open research	Secret research[2]
Very uncertain about consequences: $e_0 = 0.1, \sigma_0 = 3$			
[wide flat distribution plot, axis labeled π, origin 0]	Project team is widely uncertain about project effects.	Secret research[1]	Scrap idea
Foresee positive & negative consequences: $e_0 = 0.1, \sigma_0 = 5$			
[bimodal distribution plot, two peaks, axis labeled π, origin 0]	Some predict huge success, others predict negative consequences.	Secret research[1]	Scrap idea

Table visualizes the two pathways to secret innovation. Superscripts 1 and 2 correspond with each pathway. For pathway 1, column 1 plots different distributions for initial expectations about what affect an innovation will have for advancing the agency's goals ((π)). Treat Row 1 as the baseline. Then each subsequent row varies the average expectation e_0 of affects, and variance surrounding that expectation σ_0. Column 3 explains the research pathway in equilibrium given k_D is low. For pathway 2, we contrast how equilibrium choices vary when we move from a low to high cost (columns 3 and 4) holding constant initial expectations of success.

The second way initial ideas vary is in the standard error of $p()$. We notate it σ_0. Substantively, a high standard error could represent two things. At the individual-level (row 3), it represents an idea that is so novel there is little else to compare it to. In these cases, researchers do not know what to expect but accept that unleashing the idea on the world could have many unanticipated consequences.

At the group-level (row 4) σ_0 represents disagreement about the potential consequences of innovation. The debate surrounding autonomous weapons systems (AWS) offers a useful example. Proponents of AWS emphasize greater speed, and stealth on the battlefield, with fewer military casualties.[29] Critics point out that AWS might create greater instability and lead to more crises.[30] Before these systems are deployed, it is hard to know if they will benefit us, or cause harm.

The following expectation summarizes one pathway to research under the assumption that the political costs associated with production are low (column 3):

> **Pathway 1: Deep uncertainty.** If the political cost associated with research is low, then secret research facilitates innovation if:
>
> - R is unsure if the innovation will yield benefits or costs once deployed ($e_0 \approx 0$). If instead she was confident that it will yield benefits $e_0 >> 0$ she would pursue open research.
> - There is little pre-existing scientific research. Therefore, the researcher is not confident in her initial expectation (σ is high). If instead she was more confident that she understood the idea's effects (σ was lower), she would scrap the idea.

Why does secrecy facilitate research when researchers are deeply uncertain about the project's effects? The logic relies on two steps. In Lemma 3.1 we showed that the manager only pursues research if her expected benefits for success are sufficiently high. Deep uncertainty implies an idea could generate large positive effects or large negative effects. When D weighs these different outcomes his expectation of benefits is near 0. This is what we observe in rows, (a), (c) and (d) of Table 1. Of course, D could use research to learn more about whether the idea is viable. However, research is costly and D's expectation that pilot research will show promise is tied to his initial expectation of the innovation's effects (i.e. approximately 0).

In proposition 3.2 we showed that the researcher is also sensitive to expected benefits, but is willing to pursue research under more conditions because she can distribute the costs. As a

---

[29]https://www.nscai.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf.
[30]https://www.rand.org/blog/2020/06/the-risks-of-autonomous-weapons-systems-for-crisis.html.

result, when the costs and expected benefits are both low, the researcher is willing to pursue secret research *so long as she believes that her research will convince the manager to approve her idea.* The researcher believes that a manager is likely to be convinced by pilot research when σ_0 is high.

One reason for this is that when there is little pre-existing research, the researcher's pilot research report carries a larger weight in the manager's overall expectation of success. Another is that when projects are likely to have either extreme positive or negative consequences, then pilot testing tells us which direction the program is going to go. If the results are positive, D is confident the project will have major benefits and can accrue those by authorizing the project.

The second pathway to secret innovation relies on a trade-off between the political costs of research (c_D) and the expected consequences of deploying a new innovation (e_0). Substantively, c_D captures how sensitive the manager (and the institution at large) is to the moral and political costs associated with research.

> **Pathway 2: High stakes.** Secret research facilitates innovation when:
> - the expected benefit of innovation is high (e_0 is high); and
> - the manager's sensitivity to the political costs of research are also high (c_D is high).
>
> If the political costs of research and production were lower, we would observe open research.

The logic for the basic trade-off is simple. There are initial ideas that show enormous promise. However, the research required to pursue these ideas involves political costs. Secrecy facilities innovation when the manager is unwilling to bear the large costs of research and the costs of approval on his own. But once the research is complete, the manager will happily approve the project. In cases like this, the researcher may bear x share of the research costs knowing the manager will bear the approval costs once research is complete.

3.2.2 Predictions about patterns of innovation: Secret institutions generate important innovations that open institutions do not

In terms of aggregate patterns of innovation, what are the features of research projects and innovations we expect to see from secret versus open institutions? We find that secrecy allows organizations to research ideas that seem bizarre, morally repugnant, and likely to fail when first conceived. This leads to a straightforward expectation.

14

Expectation 1 *A larger proportion of ideas are rejected after secret research than after open research.*

We might intuit from this that secret innovation damages a nation's security in the aggregate. However, secret organizations are only willing to pursue these ideas because the potential upside is enormous. The initial idea must have a large enough chance of making a positive impact for a researcher to pursue it. If research confirms the idea is harmful, the institution scraps it early on. In the rare cases when research suggests an idea will provide benefits, these ideas are converted into innovations that change the world. This leads to a second prediction:

Expectation 2 *Secret research leads to radical innovation. For ideas put into production following secret research, the difference between their true impact and our initial expectation of their impact is large relative to ideas put into production following open research ($\pi - e_0$ is large).*

Thus, for every handful of bizarre and shameful research projects that fail—bionic cat robots, nuclear-induced tsunamis, and so forth (see Houghton, 2019)—secret institutions gives us one radical success—microchips, autonomous robots, satellites and the lithium-ion battery. With foresight, these innovations all sounded the same. With hindsight, some are radical innovations that shaped the industrial and digital revolution, and medical sciences.

3.3 Connecting the mechanism to the principal-agent literature

The basic model identified how secret innovation allowed agents to distribute costs at different stages of the innovation process so that they could pursue a wider range of novel ideas. However, the model did not adequately explore the perverse incentives that arise given uncertainty and principal-agent situations. We now introduce principal-agent problems into the model. We show that our basic logic survives, and we derive additional implications about how researchers and managers collaborate to exploit secrecy in national security institutions.

3.3.1 Monitoring

We assumed that the researcher is in total control over secret research. If researchers exploit secrecy, the manager is forced to take on a $c_D x$ cost when the program comes to light. In practice,

managers can monitor subordinates' activities by asking for details of a project. Given that the researcher's actions can force the manager to incur costs, it seems like the manager would want to monitor the researcher's activities.

This insight is at the heart of the principal-agent framework: Managers want to stop subordinates from taking actions that they would not approve of. In this literature, there is agency loss because D finds it difficult or expensive to monitor R. Thus, subordinates pursue projects because they believe monitoring is too costly for their managers, and they will therefore get away with it. However, if the manager paid no cost to monitor they always would. Anticipating certain monitoring, the researcher would always behave (Eisenhardt, 1989). This concern is relevant for our theory because R only uses secrecy because D will not approve.

We adjust the model in two ways to capture monitoring as it is commonly studied in the principal agent-framework. First, we introduce uncertainty about the cost associated with research. We start with a simplifying assumption that the cost of research is equal $c = c_R = c_D$. We then add a step at the beginning of the game where Nature selects the cost associated with research $c \sim f()$ where $f()$ is supported on the non-negative real numbers. Second, we assume that if the manager does not observe open research he has the opportunity to monitor the researcher's activities. If the manager chooses to monitor and discovers the researcher started a secret research program, he has two options: he can allow the research to continue or shut it down. If the manager allows the program to continue, the game reverts to open research (and associated payoffs). If the manager shuts the research down, the manager avoids research costs entirely, the researcher incurs c_R, and research has no effect (we do not realize θ, m).

To be clear, we explicitly assume that the manager pays no cost to monitor the researcher, and if the manager does monitor, there is a 100% chance that he learns everything the researcher knows. Indeed, this is the exact condition that the principal-agent literature suggests should drive complete monitoring. Define $\bar{c} = \lambda_0[e_0 + \theta - k_R x]$, and $\underline{c} = \lambda_0[e_0 + \theta - k_D]$.

Proposition 3.3 *The don't ask don't tell equilibrium. Suppose conditions 1 and 2 can be satisfied for some $c = c_R = c_D$, then in the model where D can perfectly monitor R, if*

$$\int_{\underline{c}}^{\bar{c}} f(c)dc < \frac{\lambda_0[e_0 + \theta - k_D]}{x} \tag{3}$$

16

the following pure strategies are a Pure Bayesian Equilibrium:

- *D does not monitor, does not approve innovation absent research. D approves open research if condition 1 is violated; and innovates following research if $e_1 > k - \theta$.*

- *Off path, if D decides to monitor he shuts down any project with a cost profile $c \geq \underline{c}$.*

- *R scraps the project if $c < \underline{c}$, R conducts open research if $c > \bar{c}$ and conducts secret research otherwise.*

- *Off path, if R openly pursues research for a cost profile $c \geq \underline{c}$, D rejects it.*

Secrecy facilitates innovation as stated in proposition 3.2 if $c \in \underline{c}, \bar{c}$.

See Appendix A.4. This result is surprising. After all, the only reason the researcher does not ask for permission is that she knows the manager will not approve. Thus, when the manager observes the researcher hiding her activities, he should suspect something bad is happening and engage in monitoring. From the researcher's perspective this is indeed what is going on: she is exploiting secrecy because she knows her manager will not approve of her controversial research program. And yet, the manager elects not to monitor. Why? The logic follows a don't-ask-don't-tell dynamic made possible by cost passing. The manager knows if he monitors he will learn the devilish details of what is happening and be forced to shut down the project, rendering a payoff of 0. However, if the manager decides not to monitor, he can reduce his costs through plausible deniability.

To be clear, some research protocols involve costs that leave the manager worse off. In this equilibrium, there are research protocols that are so controversial that the manager does worse by allowing research to continue even though he only incurs an x share of the cost. Despite this extreme preference asymmetry, the equilibrium holds because the manager expects the researcher's protocol is too controversial to approve, but not so controversial that the manager does not want the researcher to pursue it in secret.

This leads to the following empirical implications:

Expectation 3 *Don't ask don't tell:* *When managers are alerted that a researcher is engaging in a secret research project, but the researcher does not want to share the details, the manager*

17

elects not to monitor because he suspects that the program is controversial. Rather, the manager allows the researcher to continue in secret so that the manager can retain plausible deniability over controversial research practices.

Expectation 4 *Telling implies shut-down:* *If the manager ever observes the controversial details of a research program that a researcher has elected to pursue in secret, the manager will shut down the parts of the program that he observes.*

3.3.2 Trust when the researcher can fabricate her report

The analysis above emphasized that secrecy has positive effects because it provides researchers with autonomy; managers with cover from political costs; and both actors the capacity to distribute costs between them. In practice, secrecy also creates greater opportunities for R to fabricate reports or cherry-pick results of pilot research to emphasize R's good work. Managers are aware of a researcher's incentives to promote her own work. In theory, it could cause the entire secret research program to unravel. Secret research only works if the manager can trust the researcher's description of pilot research results.

To address issues of trust we adjust the model to understand if the manager can assign a researcher to a project who will pursue controversial pilot research if it is necessary, and credibly reveal the results of that pilot.[31] First, we assume that if research is conducted in secret, only the researcher observes m. Second, we assume the researcher can write any (costless) report she likes: $m_r \to \mathcal{R}$.[32] When research happens in secret, the manager only observes the report m_r. We say the research report is honest if $m_R = m$ and dishonest otherwise. Third, we search for a researcher's cost profile c_R, k_R that satisfies the revelation principle. In short, we want to know if we can find a researcher who (1) is willing to conduct secret research; (2) is willing to write an honest report no matter the outcome of her pilot; and (3) that the manager will believe. We report the technical details and game tree for the modeling technology in Appendix A.5.

---

[31] Because National security agencies extensively vet employees, it is plausible that managers have detailed information about employee backgrounds.

[32] Trivially, adding dishonesty costs makes honesty easier.

Lemma 3.4 *If condition 1 is satisfied and*

$$k_D < e_0 + \theta \tag{4}$$

then we can always find a researcher that is honest, trustworthy, willing to conduct secret research.

See Appendix A.5 for a formal statement and proof. Lemma 3.4 explains that it is possible to find a researcher who can facilitate secret innovation. But what does this researcher look like? We put the answer in terms of expectations:

Expectation 5 *Secret research only works if the institution employs unscrupulous patriots. The researcher that takes on a secret research program and will report her results credibly and honestly must be:*

- *insensitive to the political and moral issues associated with research ($k_R \to 0$), but*

- *highly sensitive to the foreign policy costs associated with deploying a project ($c_R = c_K/x$).*

The first bullet-point summarizes the condition where the researcher is willing to pay the cost to conduct controversial pilot research even if the manager is not. The second bullet-point summarizes what it takes for the researcher to honestly report pilot research. To be clear, the condition on c_R for complete revelation is a strict equality that aligns R and D's preferences at the point where D must decide between approving innovation or not. However, we can still support the credible revelation of information, with honesty in some cases and dishonesty in others, with some cost asymmetry. For example, there are cases where the researcher is less sensitive to the costs of deployment ($c_R < c_K/x$) where D is still persuaded by R's research report, and innovates if R's report is positive. In this case, it is possible R wants to innovate following pilot research, but D would not innovate if he knew the truth. In these cases, R fabricates the report. Had R sent an honest report, D would have rejected it. D is aware of this risk but trusts R anyway because the results of pilot research that generate incentives for dishonesty are unlikely relative to the results of pilot research where both agents would proceed.

In short, D will trust R even if their preferences are not perfectly aligned because R is sufficiently sensitive to the foreign policy costs associated with innovation so that R does not want to approve

19

projects that are likely to fail in most cases. This result has a secondary implication about how a researcher who has selected into secret research will behave following the outcome of pilot research:

Expectation 6 *Suppose a researcher is willing to take on a secret research project. Then, if pilot research suggests that a project will fail, the researcher will terminate the research and argue against developing the project because she believes the chance of failure is high.*

4   Research Design

We illuminated a novel mechanism for secret innovation based on cost passing (section 3.2), the pathways along which this mechanism flow, and how the mechanism interacts with principal-agent concerns (section 3.3). We also derived implications for the types of innovations secret institutions produce but open institutions do not (e.g. section 3.2.2). We want to furnish evidence that supports our core mechanism and its implications for patterns of innovation. Process tracing a causal mechanism, however, has different requirements and concerns than evaluating counterfactual claims about whether an institution would have pursued an innovation if it was open.

We chose two different case designs to balance these inferential challenges. Each case has unique historical features that provide leverage on explaining one set of empirical implications. While neither case is perfect, we believe the advantages of one offset the weaknesses of the other. Thus, taken together they help us support the logic of our theory. Since each case design is so different, we motivate each one in its own section. We briefly summarize the different approaches in Table B.1. But each section will detail its unique design and explain how it supports our theory.

5   Monitoring the Soviets

One of the United States' most pressing priorities in the early years of the Cold War was gaining better understanding of the Soviet Union's capabilities.[33] Without it, there was a heightened risk of insecurity, the possibility of arms racing, and even inadvertent war. But an aggressive and capable air defense made the prospect of overflights below a certain altitude a risky endeavor. Thus, the search for a high-flying reconnaissance aircraft was on.

---

[33]https://nsarchive2.gwu.edu/NSAEBB/NSAEBB74/U2-02.pdf.

Table 2: What we use our cases for

Program goal	Mind Control	Monitoring the Soviets
We use it to:	Process trace our causal mechanism through the secret innovation of a single idea.	Identify the features of programs that secret institutions are willing to pursue that open institutions are not. Use that to illuminate how a secret option influences patterns of observed innovations.
Historical quirks we exploit	Intelligence scandals in mid-1970s sparked a Congressional inquiry into the program 10 years after it was complete. Exposed three findings based on testimony and documents not normally revealed for top-secret CIA scientific programs: (1) Why CIA managers gave researchers the right of secrecy and selected research staff; (2) Who knew (and did not) about controversial research activities; (3) The reporting structures inside the CIA. Further, the project goes on for years, involves accidents and ends in failure. We can exploit these features to test modeling extensions.	The weakly institutionalized national security architecture circa 1950 has three implications. (1) The Air Force, an organization that could not give researchers the right to research in secret from Air Force leadership and the president, had first right of refusal on reconnaissance projects. (2) Air Force staff and potential innovators believed the Air Force was the only game in town. (3) Lurking in the background was a newly created secret institution that could explore projects and provide researchers with internal secrecy. Once this project was complete, the CIA assumed control over some surveillance functions and other new secret agencies were stood up. Thus, we cannot exploit this design for later projects.
Coding advantages	There is a clear break between researchers, who could exploit secrecy and know all project details, and managers, who did not.	We have the complete list of initial ideas submitted to the Air Force for funding and the reason they funded some and discarded others. We have extensive project plans to score each project on our model parameters.
Pathway for secret innovation	High cost of research/ high expected reward	Uncertain benefits, high variance in outcome
Weaknesses	Only one team at CIA (a secret institution) attempts one kind of research (mind control). Cannot definitively evaluate: (a) What an open institution would have done; (b) What researchers would do if research practices were less controversial.	(a) Occurs in a unique institutional period. Unclear if the logic extends to when the CIA/military employed dedicated research teams to overhead reconnaissance. (b) Potential innovators never given the option between secrecy and openness. Cannot definitively evaluate what they would have picked.

The initial effort was spearheaded by the Air Force and various affiliated organizations. One of the most notable efforts was spearheaded by the Wright Air Development Command led by Major John Seaberg. In March 1953, Seaberg settled on desired specifications for the aircraft. He wanted it to "have an optimum subsonic cruise speed at altitudes of 70,000 feet or higher over the target, carry a payload of 100 to 700 pounds of reconnaissance equipment, and have a crew of one" (Pedlow and Welzenbach, 1992, 8). Seaberg solicited proposals from a number of smaller airframe manufacturing companies. He was seemingly interested in any solution that met his specifications and believed smaller companies would take the project more seriously and move more quickly (Pedlow and Welzenbach, 1992, 8). He heard four bids:

- Fairchild Engine and Airplane Corporation proposed a single-engine aircraft, the M-195, which promised to reach a maximum altitude of 67,200 feet.

- Bell Aircraft Corporation proposed a twin-engine plane, the Model 67, or later the X-16, which promised to reach 69,500 feet.

- Glenn L. Martin Company proposed "a big-wing version of the B-57 called the Model 294, which was expected to cruise at 64,000 feet."

- Lockheed Aircraft Corporation proposed a modified, single engine aircraft that approximated sailplane, the CL-282, which promised to reach just north of 70,000 feet (Pedlow and Welzenbach, 1992, 9).

In a moment we will support our theory by examining who funds what and why. Before that, we emphasize the unique features of this case that help us validate our core counterfactual claim.

5.1 Counter-factual reasoning at this unique period in history

Our theory is built on a counter-factual claim: secret institutions pursue research that more open institutions would reject. This is difficult to validate in the modern institutional context for three reasons. First, the military and intelligence organizations employ many scientists who devise ideas on their own. When a CIA scientist conceives of a novel idea and explores it, for example, we cannot know whether the military would have rejected it. Second, scientists and engineers select into the institutions they work for. As such, we cannot know if CIA scientists are similar to military

scientists and vice versa. Finally, private companies that devise new ideas know they can pitch them to highly secret parts of the government like the CIA through classified contract mechanisms. If our theory is right, we may never observe them take ideas to the military.

A confluence of factors in this case provides a unique opportunity to test our theory. First, the companies that bid on reconnaissance aircraft all believed that the Air force was effectively the sole outlet for such pitches.[34] Interestingly, however, a relevant secret organization did exist. In July 1954, President Eisenhower tapped the President of MIT, James Killian, to head a group of scientific experts called the Technology Capabilities Panel (TCP) (Richelson, 2002, 11). Its existence was not widely known: "As with other secret panels formed by chief executives to deal with intelligence matters, Congressional input was missing from the TCP deliberations and few Congressmen knew it even existed, although many of its decisions had an immense impact on the nation's military and intelligence preparedness (5)."[35]

Project Three, one of three entities comprising the TCP, was a small group broadly focused on intelligence capabilities. It was not specifically tasked with developing proposals for overhead reconnaissance aircraft. Thus, the small and secretive Project Three members were not soliciting bids for such aircraft, and nobody expected that they would. However, the extreme secrecy that surrounded Project Three meant that they could develop research ideas in small teams that outsiders would not know about. Thus, unlike the Air Force, they exhibit the internal secrecy that our theory requires for secret innovation.

Based on this context, it is reasonable to assume that the Wright Air Development Command and any other relevant Air Force-related entity would hear all bids pertaining to overhead reconnaissance and had first right of refusal. Moreover, any project they did fund would at least be scrutinized by the broader Air Force leadership and possibly Congress. They would have also likely believed that anything they rejected would not be funded. However, as just noted, Project Three was quietly lurking in the background and ready to pick up rejected proposals if they so chose. This allows us to evaluate our counterfactual because we can observe: (1) what the open institution actually chose to accept and reject and; (2) given what the open institution rejected, what the

---

[34]Although the CIA had developed several branches to deal with scientific intelligence and research and development in the early- to mid-1950s, they did not have much experience at that time with technical collection systems. See Fischer 2001 [OTS history].

[35]https://www.nro.gov/Portals/65/documents/history/csnr/programs/docs/prog-hist-04.pdf.

secret institution chose to accept and reject.

5.2   Who funds what and why

The Air Force opted to pursue two proposals, the modified version of the B-57 from Martin which was viewed as a short-term solution and the Bell X-16 which promised better results in the medium-term. Bell was contracted to produce 28 such aircraft. At the same time, the Air Force rejected the Fairchild and Lockheed proposals. The Fairchild proposal was relegated to the dustbin of history. The Lockheed proposal was not. Lockheed took their proposal to various parts of the Air Force—including the Wright Air Development Command as well as Strategic Air Command and the Office of Development Planning—all of whom rejected it (Pedlow and Welzenbach, 1992, 11-12). Along the way, Project Three members learned of the Lockheed proposal and were immediately interested in it (Pedlow and Welzenbach, 1992, 31). As we will detail more in a moment, they undertook intense secretive research into CL-282's viability and verified that it would work. This project was later handed to the CIA as the U-2 project.

We predict that open organizations facilitate innovation when the benefits are clear (e_0 is positive), and there is not much disagreement about the likely effects (σ_0 is low); they will reject ideas that are radically new because they know little about them. Even though new ideas could have benefits, they could also cause damage. Open institutions are unlikely to take on projects like this even in the research phase (e_0 is near 0). Of these ideas, we predict that secretive institutions will pick them up as research projects if the potential outcomes vary widely (σ_0 is high). That is, there is a risk of catastrophic damage towards mission objects and enormous benefits that extend beyond what the other proposals could accomplish.[36]

This is precisely what we find. The Air force funded two safer projects that incrementally advanced the state of overflight. The modified B-57 is an obvious example. The goal was to "improv[e] the already exceptional high-altitude performance of the B-57 Canberra" (Pedlow and Welzenbach, 1992, 9). It "featured lengthened wings, accommodations for cameras and sensors, and uprated twin engines."[37] The Bell X-16 was slightly more advanced than the B-57. The modifications made to reduce weight and reach higher altitudes were far less radical than the

---

[36]That is, there is uncertainty about whether the innovation will move the U.S. closer to or further from its policy objectives.

[37]https://www.nasa.gov/sites/default/files/atoms/files/unlimited-horizons.pdf [page 1].

CL-282.[38]

The U-2 was radical by design. Senior Lockheed designers prioritized "nonstandard" elements, including "the elimination of landing gear, the disregard for military specifications, and the use of very low load factors" (Pedlow and Welzenbach, 1992, 10). Several elements of what was eventually dubbed the CL-282, and would later become the U-2, "were adapted from gliders. Thus, the wings and tail were detachable. Instead of conventional landing gear," Kelly Johnson, the lead developer, "proposed using two skis and a reinforced belly rib for landing—a common sailplane technique—and a jettisonable wheeled dolly for takeoff." As a declassified history of the U-2 puts it, "Essentially, Kelly Johnson had designed a jet-propelled glider" (Pedlow and Welzenbach, 1992, 12).

Part of Seaberg and the Wright Air Development Command's rationale for rejecting the CL-282 proposal speaks to their uncertainty about whether it would work. Seaberg pointed to its use "of the unproven General Electric J73 engine. The engineers at Wright Field considered the Pratt and Whitney J57 to be the most powerful engine available." All three of the other proposals they received from small manufacturers relied on the latter. Moreover, Seaberg and colleagues viewed "[t]he absence of conventional landing gear" on the CL-282 as a "shortcoming." Because the other proposals, including the most promising—the Bell—had "normal landing gear," they were considered "more conventional aircraft" (Pedlow and Welzenbach, 1992, 12-15).

Other Air Force commands also registered dismay at the novel features of CL-282. General Curtis LeMay, the head of Strategic Air Command, apparently "stood up halfway through the briefing, took his cigar out of his mouth, and told briefers, that if he wanted high-altitude photographs, he would put cameras in his B-36 bombers and added that he was not interested in a plane that had no wheels or guns." He called the meeting "a waste of his time" (Pedlow and Welzenbach, 1992, 12).[39]

According to the declassified history of the U-2, another driving factor in the Air Force's rejection of the CL-282 had to do with their "preference for multi-engine aircraft." This was based on familiarity and their experience with multi-engine aircraft during World War II and likely explains why they also opted for the Bell and Martin designed but rejected the Fairchild bid, which relied on a single engine. Moreover, "aerial photography experts" at the time "emphasized focal length as

---

[38]https://www.nasa.gov/sites/default/files/atoms/files/unlimited-horizons.pdf [4-5].

[39]LeMay's reaction illustrates one way that military culture imposes costs on innovators. As we argued, this makes innovation difficult in open institutions.

the primary factor in reconnaissance photography and, therefore, preferred large aircraft capable of accommodating long focal-length cameras" (Pedlow and Welzenbach, 1992, 13)

As the foregoing makes clear, the CL-282's novel design meant that many in the Air Force were skeptical about its chances of success. In terms of the model's parameters, the balance of Air Force staff thought the overall impact of the project would cause no benefit (or harm) for surveilling the Soviet Union and ultimately ensuring peace. However, some raised concerns which implied that it could have catastrophic effects: "there was the feeling shared by many Air Force officers that two engines are always better than one because, if one fails, there is a spare to get the aircraft back to base... Furthermore, a high-altitude reconnaissance aircraft deep in enemy territory would have little chance of returning if one of the engines failed, forcing the aircraft to descend" (Pedlow and Welzenbach, 1992, 13). In other words, there was concern that a single-engine plane that was missing key parts could crash inside the Soviet Union and conceivably spark a conflict.

To be sure, not everyone in the Air Force shared the view that the Bell and Martin proposals were superior to the CL-282. Trevor Gardner, Special Assistant for Research and Development, and some other officials thought it had potential. They believed "it gave promise of flying higher than the other designs and because at maximum altitude its smaller radar cross-section might make it invisible to existing Soviet radars" (Pedlow and Welzenbach, 1992, 15). Thus, if it worked, its value would be larger than the other projects.

Taken together, these divergent views support the notion that there was deep uncertainty about what CL-282 would accomplish. While some believed it was unlikely to work and therefore have no effect, others thought it could have either very negative or very positive (i.e. more positive than the other designs) effects. If the Air Force had been the only organization that could have considered the overflight proposals, one of the most important innovations of the twentieth century may never have seen the light of day (Pocock, 2000, 14).

Project Three members were themselves sensitive to the risks associated overflight over the Soviet Union.[40] But despite these risks, they pursued the project because of the enormous potential upside if the project was successful. "By the end of October [1954], the Project Three meetings had covered every aspect of the Lockheed design. The CL-282 was to be more than an airplane with a camera, it was to be an integrated intelligence-collection system that the Project Three members

---

[40]See https://history.state.gov/historicaldocuments/frus1950-55Intel/d194.

were confident could find and photograph the Soviet Union's Bison bomber fleet and, thus, resolve the growing 'bomber gap' controversy." They were also taken with the prospect that the proposal could be "the platform for a whole new generation of aerial cameras" (Pedlow and Welzenbach, 1992, 31).

Their approach to research supports our theory in two additional ways. First, they operated in secret. Land and his team "began developing it into a complete reconnaissance system," meeting in small-group settings with usually less than 10 people present. Second, they did not instantly recommend production of U-2 planes. Rather, they exploited secrecy to determine if the project was viable. Once they realized it was, they revealed what they had been doing to the CIA Director and to President Eisenhower who was extremely receptive. He "approv[ed] the development of the system, but ... stipulat[ed] that it should be handled in an unconventional way so that it would not become entangled in the bureaucracy of the Defense Department or troubled by rivalries among the services" (Pedlow and Welzenbach, 1992, 33).[41]

Interestingly, the project also helped the TCP realize that secret organizations like the CIA were well-suited to the task of overseeing radical innovations of this kind. As the TCP argued to CIA Director Allen Dulles in a memo, "this seems to us the kind of action and technique that is right for the contemporary version of the CIA; a modern and scientific way for an Agency that is always supposed to be looking, to do its looking. Quite strongly, we feel that you must always assert your first right to pioneer in scientific techniques for collecting intelligence... This present opportunity for aerial photography seems to us a fine place to start."[42]

6 Mind control

The second case traces our mechanism through one of the most controversial research initiatives of the Cold War: the CIA's search for mind control. In the late-1940s U.S. policy-makers became convinced that the Soviet Union and the People's Republic of China had mastered mind control.[43] According to Richard Helms, a longtime CIA official who would go on to become the CIA Director:

---

[41]Interestingly, the Air Force eventually comes around to accepting the proposal but does not actually abandon their X-16 program until the U-2 was operational.

[42]https://nsarchive2.gwu.edu/NSAEBB/NSAEBB74/U2-03.pdf.

[43]An important fact leading to this assessment was that cardinal Josef Mindszenty confessed to treason. U.S. officials became convinced that the Soviets had devised a way to control his mind (Marks, 1979, 21).

"There was deep concern over the issue of brainwashing... We felt that it was our responsibility not to lag behind the Russians or the Chinese in this field." (Kinzer, 2019, 54).[44]

Since the CIA was convinced the Soviets had done it, U.S. policymakers were hopeful they could unlock the mysteries for themselves (U.S. Senate, 1976, 385). To wit, policy-makers thought that mind control was "of the utmost importance... [it] could mean the difference between the survival and extinction of the United States" (Kinzer, 2019, 49). A declassified memo from the early 1950s provides a window into some of the core aims: "A. Can accurate information be obtained from willing or unwilling individuals. B. Can Agency personnel (or persons of interest to the agency) be conditioned to prevent any outside power from obtaining information from them by any known means? C. Can we obtain control of the future activities (physical and mental) of any given individual...? D. Can we prevent any outside power from gaining control of future activities (physical and mental) of agency personnel by any known means?"[45]

In 1950, the CIA conducted some ad hoc experiments under the code-name Bluebird/Artichoke.[46] Even at this early stage, projects were handled outside of the normal oversight channels. A memo to the CIA Director on April 5, 1950 stated: "In view of the extreme sensitivity of this project and its covert nature, it is deemed advisable to submit this project directly to you, rather than through the channel of the Projects Review Committee. Knowledge of this project should be restricted to the absolute minimum number of persons."[47]

Within a few years, the Agency decided to "intensify and systematize" their efforts to explore mind control. On April 13, 1953, CIA Director Allen Dulles authorized Sidney Gottlieb to establish a program called MK-ULTRA (Kinzer, 2019, 72-73). He was given even greater license to conduct experiments with virtually no oversight. Years of controversial experiments followed.

Gottlieb tested the effects of LSD on unwitting, non-volunteer subjects (U.S. Senate, 1976, 391). He also experimented on prisoners at secret prisons overseas, and on American citizens and soldiers without their consent (Kinzer, 2019). For example, under Operation Midnight Climax, prostitutes lured unsuspecting American citizens to a CIA-run safehouse in San Francisco where CIA staff secretly administered LSD and monitored them without their knowledge (Kinzer, 2019,

---

[44]For a thorough treatment of brainwashing, see https://www.cia.gov/library/readingroom/docs/CIA-RDP78-02646R000100100002-4.pdf.

[45]https://www.cia.gov/library/readingroom/docs/DOC_0000140401.pdf.

[46]https://www.cia.gov/library/readingroom/docs/CIA-RDP05S00620R000601400050-6.pdf.

[47]https://www.cia.gov/library/readingroom/docs/CIA-RDP83-01042R000800010003-1.pdf.

141-152).[48] Looking back, the Church Committee found that "the nature of the tests, their scale, and the fact that they were continued for years after the danger of surreptitious administration of LSD to unwitting individuals was known, demonstrate a fundamental disregard for the value of human life" (U.S. Senate, 1976, 386).

To be clear, when MK-ULTRA was established, Dulles and several other high-level CIA officials knew that Gottlieb's research would likely involve LSD and hypnosis (Kinzer, 2019, 54). However, as the Church Committee concluded two decades later: "Few people, even within the agencies, knew of the programs and there is no evidence that either the executive branch or Congress were ever informed of them" (U.S. Senate, 1976, 385-386). Even the senior-most managers did not know the details of the experiments. "MK-ULTRA was top secret even within the CIA. Only two officers—Gottlieb and Lashbrook—knew precisely what it was doing" (Kinzer, 2019, 154). As described below, this was by design.

In summary, several features of this case fit our high-stakes pathway for secret innovation. The case involves two primary actors: the MK-ULTRA research team (headed by Gottlieb), and CIA management (the most active member was Allen Dulles).[49] At the outset, Dulles knew that if MK-ULTRA succeed it would generate large benefits (e_0 was high). However, he also knew the necessary research would be controversial (c_i was high).[50] Starting from this position, three facts about this case match the choices that our model predicts. First, Dulles hand-selected Gottlieb to oversee MK-ULTRA. Second, Gottlieb assessed that highly controversial human subjects research was necessary to complete MK-ULTRA. Gottlieb could have discussed his research plans and results with Dulles. But instead, he chose to keep these details secret. Third, Dulles had several opportunities to learn what Gottlieb was up to but he never asked.

6.1 Why did Dulles select Gottlieb?

Gottlieb was not an obvious pick to lead MK-ULTRA. Although he had experience in government laboratories as a chemist, he did not have an intelligence background. Why did Dulles pick

---

[48]See https://www.cia.gov/library/readingroom/docs/CIA-RDP88-01315R000200230006-5.pdf and https://www.cia.gov/library/readingroom/docs/CIA-RDP88-01315R000200230022-7.pdf and https://www.cia.gov/library/readingroom/docs/CIA-RDP99-00498R000200150007-0.pdf.

[49]We emphasize Dulles because he established MK-ULTRA and appointed Gottlieb to run it. Gottlieb reported directly to Dulles. Even though Dulles did not know much, he knew more than anyone outside Gottlieb's team.

[50]To be clear, they did not know how controversial.

an intelligence outsider to lead a high stakes, and intensely secret project? In extension 3.3.2, we argued that when researchers conduct scientific tests in secret, it is easy for them to give their manager the mistaken impression that their novel idea is more effective than the research suggests. Anticipating this problem, the manager must carefully select unscrupulous patriot. That is, a researcher who is insensitive to whatever controversy it takes to complete a research program, but who shares the manager's desire to only field projects that will advance American interests (and therefore will be honest about whether the project is viable). Kinzer (2019, 47) explains that this is exactly how Dulles saw it. The CIA needed "a character steely enough to direct experiments that might challenge the conscience of other scientists, and a willingness to ignore legal niceties in the service of national security."'

The details surrounding how and why Sidney Gottlieb was chosen to head the mind control programs strongly support this expectation. According to Kinzer (2019, 50-51), "Like many Americans of his generation, he had been shaped by the trauma of World War II [which] left him with a store of pent-up patriotic fervor. His focused energy fit well with the compulsive activism and ethical elasticity that shaped the officers of the early CIA." Even though Gottlieb was new to the CIA, Kinzer notes that both Dulles and his deputy on the project, Helms, were impressed with Gottlieb's efforts to learn the tradecraft of intelligence. Pretty soon, he was heading the Chemical Division of the Technical Services Staff. When he later testified before a Senate Subcommittee about the project, Gottlieb articulated precisely the language we would expect from an unscrupulous patriot: "I would like this committee to know that I considered all this work ... to be extremely unpleasant, extremely difficult, extremely sensitive, but above all to be very urgent and important ... The feeling that we had was that there was a real possibility that potential enemies, those enemies that were showing specific aggressive intentions at that time, possessed capabilities in this field that we know nothing about, and the possession of those capabilities... combined with our own ignorance about it, seemed to us to pose a threat of the magnitude of national survival" (Kinzer, 2019, 238).

Of course, Gottlieb faces incentives to cast himself as patriotic during an inquiry into his conduct. However, his behavior before in the final years of the project also reflects actions that are consistent with the patriotic motives we describe. In our theory, the patriotic researcher only pursues her project because she believes the science is viable. If ever she learns that her research will not lead to an innovation that advances US national security interests, she will quit even if no

one stops her.

Consistent with this logic, a major reason why MK-ULTRA was eventually shut down after nearly a decade of experimentation, was that Gottlieb eventually realized that "On the scientific side, it has become very clear that these materials and techniques are too unpredictable in their effect on individual human beings, under specific circumstances, to be operationally useful" (Kinzer, 2019, 198). During his Senate testimony in 1977, he "publicly asserted the conclusion he had reported to his CIA superiors when he ended MK-ULTRA more than a decade earlier" which was that "there is no such thing as mind control" (Kinzer, 2019, 238). A researcher who was motivated by pride would not publicly declare their work a failure.

6.2   Why did Gottlieb choose secrecy?

If our theory is correct, then Gottlieb operated in secrecy because he knew that even his managers at the CIA would refuse to let him continue the experiments if they figured out what he was up to. Unfortunately, Gottlieb never explained why he kept these details from Dulles. But a closer look at his actions are consistent with our logic.[51]

First, Gottlieb did file reports about MK-ULTRA's progress (U.S. Senate, 1976, 401). However, "there were no attempts to secure approval for the most controversial aspects of these programs from the executive branch or Congress. The nature and extent of the programs were closely held secrets; even DCI McCone was not briefed on all the details of the program involving the surreptitious administration of LSD until 1963" (U.S. Senate, 1976, 394).[52] That Gottlieb selectively left out the highly controversial details supports our theory. Second, when Congress got wind of MK-ULTRA and asked to review the program files, Gottlieb destroyed them instead of handing them over to his superiors (U.S. Senate, 1976, 389,406).

Finally, several CIA managers have stated that they would have stopped MK-ULTRA if they had known about it. DCI Stansfield Turner stated that "It is totally abhorrent to me to think of using a human being as a guinea pig and in any way jeopardizing his health, no matter how great the cause... I am not here to pass judgment on my predecessors, but I can assure you that this is totally beyond the pale of my contemplation of activities that the CIA or any other of our

---

[51]The U2 case provides more comprehensive evidence of this step.

[52]There is no evidence that DCI McCone or anyone else was briefed on the issues of subject consent, or secret prisons.

intelligence agencies should undertake" (Kinzer, 2019, 234).

6.3   Managers built the system so they were in the dark

Our theory suggests that CIA managers are likely to embrace ignorance in cases like this because of the logic of cost passing. The managers know that if they do not investigate, they will incur a small cost as an ignorant by-stander. But they may accrue a large gain from a successful innovation. If they investigate, they are faced with the choice of incurring a large cost or shutting down the program altogether. Under broad conditions, they prefer to remain ignorant.

Several features of this logic play out in this case. First, Dulles made sure to minimize his exposure to MK-ULTRA details from the outset. When he initially authorized the project in 1953, the $300,000 he set aside was "not subject to financial controls" and researchers had "permission to launch research and conduct experiments at will, 'without the signing of the usual contracts or other written agreements' " (Kinzer, 2019, 73). According to one account, "Dulles ordered the Agency's bookkeepers to pay the costs blindly on the signatures of Sid Gottlieb and Willis Gibbons, a former U.S. Rubber executive who headed TSS" (Marks, 1979, 57).

Second, Dulles and other CIA managers went to extraordinary lengths to avoid looking into MK-ULTRA. The most extreme example involved a civilian employee of the Army, Frank Olson, who was unwittingly given LSD and purportedly jumped out of a hotel window to his death in the weeks afterwards. The internal investigation that followed accused the Technical Services Staff of "fail[ing] to observe normal and reasonable precautions." In response, CIA Director Dulles wrote a letter to Gottlieb "criticizing him for 'poor judgment... in authorizing the use of this drug on such an unwitting basis and without proximate medical safeguards' " (U.S. Senate, 1976, 398). Ultimately, however, these were not formal reprimands, had no effect on advancement, and did not lead to a termination of the experiments. Shockingly, even after investigators uncovered wrongdoing in the narrow experiments related to Olson, they did not expand their audit to the MK-Ultra program broadly.[53]

Third, when MK-ULTRA was eventually made public, the costs were distributed in accordance with our theory. As the most senior scientist who knew the complete details, Gottlieb was hauled

---

[53]According to one account, a senior CIA official cautioned that a formal reprimand "would hinder 'the spirit of initiative and enthusiasm so necessary in our work' " (Marks, 1979, 84).

before Congress to testify. Years later, he was implicated in a variety of lawsuits of families of victims of MK-Ultra. Most important for our purposes, "since Richard Helms was not alleged to have been directly involved in the drugging, he could not be prosecuted—but ... the case against Gottlieb could proceed" (Kinzer, 2019, 256-257). Gottlieb successfully secured a postponement of the trial, which was originally scheduled to begin on January 3, 1999. He died on March 7, 1999 (Kinzer, 2019, 257-259).

7   Conclusion

We argued that secretive national security institutions are more innovative *because* they are secret. Secrecy allows an enterprising researcher to pursue ideas that are so bizarre, morally controversial, and unlikely to work ex-ante that their manager would deny them. But if pilot research confirms the researcher's intuition, she can convert it into an innovation. These ideas reflect some of the most important innovations in the twentieth century.

Our theory complicates recent research that sees secrecy as a threat to national security; and even democracy (Carson, 2018; Carnegie and Carson, 2018; Carnegie, 2021; Colaresi, 2014). Based partially on this research, pundits have embraced greater transparency, information sharing, and oversight. For example, unauthorized media leaks about sensitive policy and national security discussion have become commonplace (cf Jervis, 2010). While the oversight benefits are known, our paper exposes long-run institutional costs that we may only start to understand decades from now.

These complications are important for the resurgence of great power competition. Consider debates about how the U.S. manages China's rise. Many argue that more national security innovation is needed to maintain an edge.[54] At the same time, many also argue that the U.S. must adhere to its democratic principles—which include oversight and openness—to maintain its edge over China.[55] Our theory suggests that these two strategies are often in tension. The return of great power competition has also created pressure to classify more innovation and punish leakers. Our theory suggests that these choices may promote novel national security innovations. Indeed,

---

[54]https://thehill.com/opinion/technology/466705-why-america-must-boldly-win-the-technological-race-against-china

[55]https://www.washingtonpost.com/opinions/2020/06/03/way-defeat-china-is-be-true-ourselves/

senior officials have noted unanticipated benefits. As Former Vice Chairman of the Joint Chiefs of Staff, Gen. John Hyten, argued: "when 'you're in the black, nobody's checking your homework' ... [which] means programs are free to have setbacks without risking cancelation."[56] Our theory highlights how the U.S. can properly harness secrecy to maintain our innovation edge.

Our theory also cautions the intelligence community against relying too heavily on venture capital initiatives like In-Q-Tel,[57] or contracting models like DISA.[58] While valuable, in-house research at the CIA Directorates of Digital Innovation, and Science and Technology; and at the NSA and NRO are essential for national security innovations.

Finally, our theory has important private-sector implications in an age of socially responsibility. Firms are increasingly asked to choose between safer, less controversial research bets, or boycotts and shareholder protests. This may explain why the most innovative entrepreneurs, such as Elon Musk and Peter Thiel, who have collectively founded seven, one-billion dollar companies are those willing to embrace political controversy.

---

[56]https://www.militarytimes.com/news/pentagon-congress/2021/10/28/risk-aversion-and-secrecy-are-costing-us-its-military-advantage-no-2-general-says/?utm$_source$ $=$ $Sailthruutm_medium$ $=$ $emailutm_campaign$ $=$ $EBB\%2010.29.2021utm_term = Editorial\%20 - \%20Early\%20Bird\%20Brief.$
[57]https://www.iqt.org/about-iqt/.
[58]https://storefront.disa.mil/kinetic/disa/service-catalog/category/cyber-security

References

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith and Peter Howitt. 2005. "Competition and Innovation: An Inverted-U Relationship*." *Quarterly Journal of Economics* 120(2):701–728.

Andrew, Christopher. 1995. "American Presidents and their Intelligence Communities." *Intelligence and National Security* 10(4):95–112.

Bateman, Aaron. 2020. "Technological Wonder and Strategic Vulnerability: Satellite Reconnaissance and American National Security during the Cold War." *International Journal of Intelligence and CounterIntelligence* 33(2):328–353.

Biddle, Stephen, Julia Macdonald and Ryan Baker. 2018. "Small footprint, small payoff: The military effectiveness of security force assistance." *Journal of Strategic Studies* 41(1-2):89–142.

Black, Sandra E. and Lisa M. Lynch. 2004. "What's Driving the New Economy?: The Benefits of Workplace Innovation." *The Economic Journal* 114(493):F97–F116.

Boushey, Graeme. 2016. Targeted for diffusion? How the use and acceptance of stereotypes shape the diffusion of criminal justice policy innovations in the American States. In *American Political Science Review*. Vol. 110 Cambridge University Press pp. 198–214.

Bresnahan, T. F., E. Brynjolfsson and L. M. Hitt. 2002. "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence." *The Quarterly Journal of Economics* 117(1):339–376.

Breznitz, Dan and Darius Ornston. 2013. "The Revolutionary Power of Peripheral Agencies." *Comparative Political Studies* 46(10):1219–1245.

Caillier, James G. 2017. "Public Service Motivation and Decisions to Report Wrongdoing in U.S. Federal Agencies: Is This Relationship Mediated by the Seriousness of the Wrongdoing." *American Review of Public Administration* 47(7):810–825.

Cain, Bruce E. 2014. *Democracy More or Less*. New York: Cambridge University Press.

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. 2018. *Secret wars: Covert conflict in international politics.* Princeton: Princeton University Press.

Coe, Andrew and Jane Vaynman. 2019. "Why Arms Control Is So Rare." *American Political Science Review* pp. 1–14.

Colaresi, Michael P. 2014. *Democracy declassified: the secrecy dilemma in liberal states.* Oxford ; New York: Oxford University Press.

Di Lonardo, Livio, Jessica S. Sun and Scott A. Tyson. 2020. "Autocratic Stability in the Shadow of Foreign Threats." *American Political Science Review* 114(4):1247–1265.

Doel, Robert E. and Allan A. Needell. 1997. "Science, Scientists, and the CIA: Balancing International Ideals, National Needs, and Professional Opportunities." *Intelligence and National Security* 12(1):59–81.

Dolnik, Adam. 2007. *Understanding Terrorist Innovation: Technology, Tactics, and Global Trends.* London and New York: Routledge.

Downs, George W. and David M. Rocke. 1994. "Conflict, Agency, and Gambling for Resurrection: The Principal-Agent Problem Goes to War." *American Journal of Political Science* 38(2):362.

Dragu, Tiberiu and Yonatan Lupu. 2021. "Digital Authoritarianism and the Future of Human Rights." *International Organization* 75(4):991–1017.

Eisenhardt, Kathleen M. 1989. "Agency Theory: An Assessment and Review." *The Academy of Management Review* 14(1):57.

Farrell, Theo G. and Terry Terriff. 2002. *The sources of military change: Culture, politics, technology.* Boulder, CO: Lynne Rienner.

Foarta, Dana and Takuo Sugaya. 2021. "Wait-and-See or Step In? Dynamics of Interventions."
American Economic Journal: Microeconomics 13(1):399–425.

Foley, Robert T. 2012. "A case study in horizontal military innovation: The German Army, 1916–
1918." *Journal of Strategic Studies* 35(6):799–827.

Frant, Howard, Frances Stokes Berry and William D. Berry. 1991. "Specifying a Model of State
Policy Innovation." *American Political Science Review* 85(2):571–579.

Freeman, Richard B. 2015. "Immigration, international collaboration, and innovation: Science and
technology policy in the global economy." *Innovation Policy and the Economy* 15(1):153–175.

Gohdes, Anita R. 2020. "Repression technology: Internet accessibility and state violence." *American Journal of Political Science* 64(3):488–503.

Gray, Virginia. 1973. "Innovation in the States: A Diffusion Study." *American Political Science
Review* 67(4):1174–1185.

Griffin, Stuart. 2017. "Military Innovation Studies: Multidisciplinary or Lacking Discipline?" *Journal of Strategic Studies* 40(1-2):196–224.

Grissom, Adam. 2006. "The future of military innovation studies." *Journal of strategic studies*
29(5):905–934.

Hawkins, D G, D A Lake, D L Nielson and M J Tierney. 2006. *Delegation and Agency in International Organizations*. Political Economy of Institutions and Decisions Cambridge University
Press.

Hirsch, Alexander V. 2016. "Experimentation and persuasion in political organizations." *American
Political Science Review* 110(1):68–84.

Hopkins, Robert S. 1996. "An Expanded Understanding of Eisenhower, American Policy and
Overflights." *Intelligence and National Security* 11(2):332–334.

Horowitz, Michael. 2010a. *The diffusion of military power : causes and consequences for international politics*. Princeton University Press.

Horowitz, Michael C. 2010*b*. "Nonstate Actors and the Diffusion of Innovations: The Case of Suicide Terrorism." *International Organization* 64(1):33–64.

Horowitz, Michael C. 2016. "Public Opinion and the Politics of the Killer Robots Debate." *Research & Politics* 3(1):1–8.

Horowitz, Michael C. 2020. "Do Emerging Military Technologies Matter for International Politics?" *Annual Review of Political Science* 23:385–400.

Houghton, Vince. 2019. *Nuking the moon: And other intelligence schemes and military plots left on the drawing board*. New York: Penguin Books.

Jervis, R. 2010. *Why Intelligence Fails: Lessons from the Iranian Revolution and the Iraq War*. Cornell Studies in Security Affairs Cornell University Press.

Jones, D. C., P. Kalmi and A. Kauhanen. 2006. "Human Resource Management Policies and Productivity: New Evidence from An Econometric Case Study." *Oxford Review of Economic Policy* 22(4):526–538.

Joseph, Michael F., Michael Poznansky and William Spaniel. 2022. "Shooting the Messenger: The Challenge of National Security Whistleblowing." *Journal of Politics* Forthcoming.

Jungdahl, Adam M and Julia M Macdonald. 2015. "Innovation inhibitors in war: Overcoming obstacles in the pursuit of military effectiveness." *Journal of Strategic Studies* 38(4):467–499.

King, Nigel. 1990. "Innovation at work: The research literature.".

Kinzer, Stephen. 2019. *Poisoner in chief: Sidney Gottlieb and the CIA search for mind control*. New York: Henry Holt and Company.

Kollars, Nina. 2017. "Genius and mastery in military innovation." *Survival* 59(2):125–138.

Kopel, Michael and Christian Riegler. 2006. "Delegation in an R&D Game with Spillovers." *SSRN Electronic Journal* .

Kuo, Kendrick. 2020. "Military Innovation and Technological Determinism: British and US Ways of Carrier Warfare, 1919–1945." *Journal of Global Security Studies* .

Lai, Edwin L.-C., Raymond Riezman and Ping Wang. 2009. "Outsourcing of innovation." *Economic Theory* 38(3):485–515.

Laursen, K. 2003. "New human resource management practices, complementarities and the impact on innovation performance." *Cambridge Journal of Economics* 27(2):243–263.

Lee, Caitlin. 2019. "The Role of Culture in Military Innovation Studies: Lessons Learned from the US Air Force's Adoptino of the Predator Drone, 1993-1997." *Journal of Strategies Studies* pp. 1–35.

Macdonald, Julia M. 2015. "Eisenhower's Scientists: Policy Entrepreneurs and the Test-Ban Debate 1954–1958." *Foreign Policy Analysis* 11(1):1–21.

Marks, John. 1979. *The search for the "Manchurian candidate": The CIA and mind control.* London: Allen Lane.

McCarthy, David S. 2013. "'The Sun Never Sets on the Activities of the CIA': Project Resistance at William and Mary." *Intelligence and National Security* 28(5):611–633.

Miller, Gary J. 2005. "THE POLITICAL EVOLUTION OF PRINCIPAL-AGENT MODELS." *Annual Review of Political Science* 8(1):203–225.

Mintrom, Michael. 1997. "Policy entrepreneurs and the diffusion of innovation." *American journal of political science* pp. 738–770.

Moghadam, Assaf. 2013. "How al Qaeda innovates." *Security Studies* 22(3):466–497.

Pedlow, Gregory W. and Donald E. Welzenbach. 1992. *The Central Intelligence Agency and Overhead Reconnaissance: The U-2 and Oxcart Programs, 1954-1974.* Washington, D.C.: History Staff, Central Intelligence Agency.

Perkoski, Evan. 2019. *Terrorist technological innovation.* Oxford: Oxford University Press.

Pocock, Chris. 2000. *The U-2 Spyplane: Toward the Unknown.* Atglen, PA: Schiffer Publishing.

Posen, Barry. 1984. *The sources of military doctrine: France, Britain, and Germany between the world wars.* Ithaca: Cornell University Press.

39

Price, John F. 2014. "US Military Innovation: Fostering Creativity in a Culture of Compliance." *Air & Space Power Journal* 43(Sep.-Oct.):128–134.

Richelson, Jeffrey T. 2002. *The wizards of langley: Inside the CIA's Directorate of Science and Technology.* Boulder, CO: Westview Press.

Rosen, Stephen Peter. 1988. "New ways of war: understanding military innovation." *International security* 13(1):134–168.

Rotemberg, Julio J and Garth Saloner. 1994. "Benefits of Narrow Business Strategies." *The American Economic Review* 84(5):1330–1349.

Sagar, Rahul. 2013. *Secrets and Leaks: The Dilemma of State Secrecy.* Princeton: Lynne Rienner.

Sechser, Todd S., Neil Narang and Caitlin Talmadge. 2019. "Emerging technologies and strategic stability in peacetime, crisis, and war." *Journal of Strategic Studies* 42(6):727–735.

Taylor, Mark Zachary. 2016. *The Politics of Innovation.* Oxford University Press.

Tushman, Michael L. and Philip Anderson. 1986. "Technological Discontinuities and Organizational Environments." *Administrative Science Quarterly* 31(3):439.

U.S. Senate. 1976. *Final report of the select committee to study governmental operations with respect to intelligence activities.* Washington, D.C.: U.S. Government Printing Office.

Vaynman, Jane. 2022. "Better Monitoring and Better Spying: The Effects of Emerging Technology on Cooperation.".

Wallace, Robert, H. Keith Melton and Henry R. Schlesinger. 2009. *Spycraft: The secret history of the CIA's spytechs, from communism to Al-Qaeda.* New York: Plume.

West, Michael A. and Neil R. Anderson. 1996. "Innovation in top management teams." *Journal of Applied Psychology* 81(6):680–693.

Winchester, Simon. 2019. *Exactly : How Precision Engineers Created the Modern World.* London: HarperCollins.

Zhang, Baobao, Markus Anderljung, Lauren Kahn, Noemi Dreksler, Michael C. Horowitz and Allan Dafoe. 2021. "Ethics and Governance of Artificial Intelligence: Evidence from a Survey of Machine Learning Researchers." *Journal of Artificial Intelligence Research* 71:591–666–591–666.

Zoghi, Cindy, Robert D. Mohr and Peter B. Meyer. 2010. "Workplace organization and innovation." *Canadian Journal of Economics/Revue canadienne d'économique* 43(2):622–639.

Appendix

Table of Contents

A   Formal Appendix

A.1   Lemma 3.1: When open institutions do not innovate

In the open institution, there are three strategy profiles that can lead to innovation. In the first R pursues an idea, D does not approve research, then D selects innovation absent research. If this was an equilibrium, then D could not profitably deviate to rejecting innovation at the final decision-node. But D prefers to deviate to rejecting innovation if $e_0 - k_D < 0 \equiv e_0 < k_D$. This cannot be satisfied if condition 1 is.

In the second pathway, R pursues an idea, D approves research, then D approves innovation after research. Off the path, if D does not research D does not innovate. In this pathway, D approves innovation at the final on-path node if $e_1 + \theta - c_D - k_D > -c_D \equiv e_1 > k_D - \theta$. Working backwards, D's expected utility for authorizing research is incentives to research are $pr(e_1 + \theta > k_D)(e_0 - k_D) - c_D$. If instead, D does not research he gets 0. This solves for the threshold of condition 1.

In the third pathway, R pursues an idea, D approves research, then D approves innovation after research. Off the path, if D does not research D approves innovation. Working backwards, D's expected utility for authorizing research is incentives to research are $pr(e_1 + \theta > k_D)(e_0 - k_D) - c_D$. If instead, D does not research he gets $e_0 - k_D > 0$. For the equilibrium to hold together, D cannot profitably deviate to no research and then innovation. This is only true if $0 < e_0 - k_D < \lambda_0(e_0 + \theta - k_D) - c_D)$. But these conditions cannot be jointly satisfied if condition 1 is.

This completes the proof.

A.2   Proposition 3.2: Secrecy facilitates innovation

We re-state the equilibrium strategies. R's strategy is to engage in secret research. D's on-path strategy is to approve a innovation if secret research yields a signal that shifts D's posterior belief $e_1 \geq k - \theta$, and reject the project otherwise. Off the path, if R asks for permission to conduct research, D does not approve research and does not approve innovation.

Notice that if R does ask for permission, we are in a sub-game that exactly reflects the open institution. It follows that D's off-path strategy to reject research and reject innovation is supported if Condition 1 is. This yields a pay-off of $0, 0$. Similarly, if R rejects an idea at the first node, pay-offs are 0 for both players.

Turning to on-path actions, in the final node, D approve an innovation if $e_1 + \theta - k_D - c_D x \geq -c_D x$ is and not otherwise. Notice that this reduces to the condition we use to derive λ_0.

Working backwards, consider R's on-path decision to engage in secret research. R's value from secret research is: $(1 - \lambda_0)0 + \lambda_0(e_0 - k_R x) - c_R$. R prefers this to all her other options (which each yield 0) when equilibrium condition 2 is satisfied. It follows that R cannot profit from deviating to open research, or rejecting under the two conditions stated in the equilibrium.

1

A.3 Two pathways to innovation

In this section we explain how we derive our empirical implications from the main model. The basic idea is to re-arrange the equilibrium conditions 1, 2 with regard to our main parameters.

The first pathway describes different features of $p()$. Our first bullet point relates to e_0, the expected value of p. Define a second distribution p_α as p that is identical to p but shifts the mean by $\alpha \in \mathcal{R}$. That is, for an arbitrary input a, $p(a) = p_\alpha(a + \alpha)$. Notice the standard deviation of p_α is σ_p and the expected value is $e_0 + \alpha$.

We can re-write the first bullet point in pathway 1 as follows. Suppose $\alpha = 0$, and otherwise we take a list of parameters that meet the conditions for secret innovation outlined in proposition 3.2. There exists a $\bar{\alpha}$ large enough so that for any $\alpha > \bar{\alpha}$ the right-most inequality of condition 2 is no longer satisfied and also $|e_0 - 0| < |e_0 + \alpha - 0|$. Further, there exists a $\underline{\alpha}$ small enough so that for any $\alpha < \underline{\alpha}$ the left-most inequality of condition 2 is no longer satisfied and also $|e_0 - 0| < |e_0 + \alpha - 0|$.

Starting with the right-most inequality, notice that $\lambda_0 - \frac{c_D}{e_0 + \theta - k_D} \to -\infty$ as $\alpha \to \infty$ and also that $\lambda_0 - \frac{c_R}{e_0 + \theta - x k_D} \to \infty$ as $\alpha \to -\infty$.

It follows that for every set of parameters for which we can observe secret innovation, that if we move the agents' expectation that research will succeed far enough away from 0 in either direction that we will break the result. To be clear, there are cases where we cannot support secret innovation for $e_0 = 0$ but can for other parameters. The point is that the parameters for which we can support it must be sufficiently close to $e_0 \approx 0$.

Our second bullet point relates to the standard error of p, σ_0. Define a second distribution p_β as p with a resolution parameter $\beta > 0$. We also define the standard error as σ_β. More precisely, for an arbitrary input a, $p(e_0 + a) = p_\beta(e_0 + \beta a)$. Notice that the expected value of p_β is equal to e_0. Also, for $\beta < 1, \sigma_\beta < \sigma_0$, and for $\beta > 1, \sigma_\beta > \sigma_0$.

We can re-write the second bullet point as follows. Suppose a list of parameters where condition 2 is satisfied and replace p with p_β, $\beta = 1$. There exists a $\underline{\beta}$ small enough so that for any $\beta < \underline{\beta}$ the left-most inequality of condition 2 is no longer satisfied. This rests on λ_0. R's benefit from secret research comes when it provides a sufficiently positive message that D will approve. That is we need an m that satisfies.

A.4 Proposition 3.3: Monitoring and principle agent dynamics

Under the assumption that D will not monitor, $\underline{c} = \lambda_0[e_0 + \theta - k_D]$ represents the level of cost associated with research that leaves D indifferent between approving open research and no research: $\lambda_0[e_0 + \theta - k_D] - c > 0$. When $c < \underline{c}$, D will approve research into the project if R asks. We've shown that R strictly prefers open research if D will approve it.

Under the assumption that D will not monitor, $\bar{c} = \lambda_0[e_0 + \theta - k_R x]$ represents the level of cost associated with research that leaves R indifferent between secret research, given R's belief that research will be pivotal for D's choice, and no research: $\lambda_0[e_0 + \theta - k_R x] - c > 0$. When $c > \bar{c}$, R prefers to scrap the project yielding a benefit of 0 for both players.

Working backwards, if D does not observe research, he knows that either R has scrapped the

project, or that R is pursuing research in secret. Thus, D prefers not to monitor, rather than to monitor (yielding a payoff o 0) if:

$$pr(c \in \bar{c}, \infty)0 + pr(c \in \underline{c}, \bar{c})[\lambda_0(e_0 + \theta - k_D) + x \int_{\underline{c}}^{\bar{c}} f(c)dc] > 0 \tag{5}$$

This solves for condition 3 in the manuscript as desired.

The Remark describes D's expected value for failing to monitor at the monitoring choice node, given every realizable value of $c \in \underline{c}, \bar{c}$. We compute it by subbing in c for $\int_{\underline{c}}^{\bar{c}} f(c)dc$ in equation 5 and solving for c. The Remark simply highlights that there are realizations where D's expected value at that node for allowing research to proceed on path is negative.

A.5   Lemma 3.4: Researcher can fabricate her report

The extension requires us to re-define the agent's beliefs about π now that they can diverge. We continue to define e_0 as both agents' prior expected value of $\pi|p$, and e_1 as both player's common post-open research expected value of π. We now define R's post-secret research expected value of π as: $e_R = e_1$ and D's post-secret research expected value of π as $e_D|s^R(m_R), p$. The only difference is that e_D depends on R's research report, and R's strategy. A strategy for R is $s^R(a_1 \in (sp, or, sr), m_R|a_1 = sr)$. A strategy for D is a double $s^D(b_1 \in (r, nr)|a_1 = or, b_2 \in (a, na)|a_1 \neq sp)$.

We ask, can we find a cost profile for a researcher c_R, k_R in which we can support honest, credible research that facilitates innovation. That is, we want to find an equilibrium that has the following properties. Research would not happen in the open institution (condition 1 is satisfied). R conducts research in secret, then for any pilot research that R observes, R sends an honest message: $m_r = m$. D believes R's message such that $e_D = e_R$. D's decision to research is conditional on R's message. Off the path, if R researchers in the open, D rejects.

The first thing to note is that there is a single cut-point of D's post-research beliefs that determines whether he approves or rejects an innovation. Assume an equilibrium in which R's research report is credible. Define a post-research belief e_D^* based on a credible research report m^* that leaves D indifferent between approval or rejection:

$$e_D^*|m^* + \theta - k_D - xc_D = -xc_D \equiv e_1^*|m^* = k_D - \theta$$

It follows that if $m > m^*$, that D will approve innovation following R's credible message and not otherwise.

Second, R's expected value from innovation following secret research is strictly increasing in m so long as D approves the project. Third, in an equilibrium characterized by honest, credible reporting, R's expected value from sending a message $m \geq m^*$ is $e_R - xk_R - c_R$; and 0 from sending $m < m^*$.

Since R's report is costless to write, and there is no restriction on m_R, R can clearly influence D's choice by deviating to a dishonest message. For R's research report to be credible, it must be that R cannot profitably deviate from honesty. The three points imply this is the case if after observing secret research R is indifferent between innovation and not given $m = m^*$. This is the case when

3

$k_R = k_D/x$ as stated in the manuscript.

We now need to check that R is willing to conduct secretive research given R's sensitivity to deployment costs. We can if: $\lambda_0[e_0 + \theta - k_R x] - c_R > 0$. Subbing in $k_R = k_D/x$ we get: $c_R < \lambda_0[e_0 + \theta - k_D]$. Since c_R must only be non-negative, we take it at the limit $c_R = 0$. Thus, we can find a researcher who is willing to research if: $e_0 + \theta = k_D$ as desired.

B   Research Design

Our theory illuminated a novel mechanism for secret innovation based on cost passing (section 3.2), the pathways along which this mechanism flow, and how the mechanism interacts with principal-agent concerns (section 3.3). We also derived implications for the types of innovations secret institutions produce but open institutions do not (e.g. section 3.2.2). We want to furnish evidence that supports our core mechanism and its implications for patterns of innovation. Process tracing a causal mechanism, however, has different requirements and concerns than evaluating counterfactual claims about whether an institution would have pursued an innovation if it was open.

We chose two different case designs to balance these inferential challenges. Each case has unique historical features that provide leverage on explaining one set of empirical implications. First, the CIA's search for mind control has unique features that allow us to support our causal mechanism. Second, the weakly institutionalized environment that surrounded the development of the U-2 allows us to support our counterfactual claims about the types of ideas secret institutions pursue.While neither case is perfect, we believe the advantages of one offset the weaknesses of the other. Thus, taken together they help us support our theory.

Since each case design is so different, we motivate each one in its own section. The MK-Ultra Case, which we use primarily for process tracing, appears in the main body. The U-2 case, which we use to illustrate the counterfactual logic of our argument, appears below.

We briefly summarize the different approaches in Table B.1.

C   Considering the Counterfactual: Monitoring the Soviets

One of the United States' most pressing priorities in the early years of the Cold War was gaining better understanding of the Soviet Union's capabilities.[59] Without it, there was a heightened risk of insecurity, the possibility of arms racing, and even inadvertent war. But an aggressive and capable air defense made the prospect of overflights below a certain altitude a risky endeavor. Thus, the search for a high-flying reconnaissance aircraft was on.

The initial effort was spearheaded by the Air Force and various affiliated organizations. One of the most notable efforts was spearheaded by the Wright Air Development Command led by Major John Seaberg. In March 1953, Seaberg settled on desired specifications for the aircraft. He wanted it to "have an optimum subsonic cruise speed at altitudes of 70,000 feet or higher over the target, carry a payload of 100 to 700 pounds of reconnaissance equipment, and have a crew of one" (Pedlow and Welzenbach, 1992, 8). Seaberg solicited proposals from a number of smaller airframe manufacturing

---

[59]https://nsarchive2.gwu.edu/NSAEBB/NSAEBB74/U2-02.pdf.

Table B.1: What we use our cases for

Program goal	Mind Control	Monitoring the Soviets
We use it to:	Process trace our causal mechanism through the secret innovation of a single idea.	Identify the features of programs that secret institutions are willing to pursue that open institutions are not. Use that to illuminate how a secret option influences patterns of observed innovations.
Historical quirks we exploit	Intelligence scandals in mid-1970s sparked a Congressional inquiry into the program 10 years after it was complete. Exposed three findings based on testimony and documents not normally revealed for top-secret CIA scientific programs: (1) Why CIA managers gave researchers the right of secrecy and selected research staff; (2) Who knew (and did not) about controversial research activities; (3) The reporting structures inside the CIA. Further, the project goes on for years, involves accidents and ends in failure. We can exploit these features to test modeling extensions.	The weakly institutionalized national security architecture circa 1950 has three implications. (1) The Air Force, an organization that could not give researchers the right to research in secret from Air Force leadership and the president, had first right of refusal on reconnaissance projects. (2) Air Force staff and potential innovators believed the Air Force was the only game in town. (3) Lurking in the background was a newly created secret institution that could explore projects and provide researchers with internal secrecy. Once this project was complete, the CIA assumed control over some surveillance functions and other new secret agencies were stood up. Thus, we cannot exploit this design for later projects.
Coding advantages	There is a clear break between researchers, who could exploit secrecy and know all project details, and managers, who did not.	We have the complete list of initial ideas submitted to the Air Force for funding and the reason they funded some and discarded others. We have extensive project plans to score each project on our model parameters.
Pathway for secret innovation	High cost of research/ high expected reward	Uncertain benefits, high variance in outcome
Weaknesses	Only one team at CIA (a secret institution) attempts one kind of research (mind control). Cannot definitively evaluate: (a) What an open institution would have done; (b) What researchers would do if research practices were less controversial.	(a) Occurs in a unique institutional period. Unclear if the logic extends to when the CIA/military employed dedicated research teams to overhead reconnaissance. (b) Potential innovators never given the option between secrecy and openness. Cannot definitively evaluate what they would have picked.

companies. He was seemingly interested in any solution that met his specifications and believed smaller companies would take the project more seriously and move more quickly (Pedlow and Welzenbach, 1992, 8). He heard four bids:

- Fairchild Engine and Airplane Corporation proposed a single-engine aircraft, the M-195, which promised to reach a maximum altitude of 67,200 feet.

- Bell Aircraft Corporation proposed a twin-engine plane, the Model 67, or later the X-16, which promised to reach 69,500 feet.

- Glenn L. Martin Company proposed "a big-wing version of the B-57 called the Model 294, which was expected to cruise at 64,000 feet."

- Lockheed Aircraft Corporation proposed a modified, single engine aircraft that approximated sailplane, the CL-282, which promised to reach just north of 70,000 feet (Pedlow and Welzenbach, 1992, 9).

In a moment we will support our theory by examining who funds what and why. Before that, we emphasize the unique features of this case that help us validate our core counterfactual claim.

C.1 Counter-factual reasoning at this unique period in history

Our theory is built on a counter-factual claim: secret institutions pursue research that more open institutions would reject. This is difficult to validate in the modern institutional context for three reasons. First, the military and intelligence organizations employ many scientists who devise ideas on their own. When a CIA scientist conceives of a novel idea and explores it, for example, we cannot know whether the military would have rejected it. Second, scientists and engineers select into the institutions they work for. As such, we cannot know if CIA scientists are similar to military scientists and vice versa. Finally, private companies that devise new ideas know they can pitch them to highly secret parts of the government like the CIA through classified contract mechanisms. If our theory is right, we may never observe them take ideas to the military.

A confluence of factors in this case provides a unique opportunity to test our theory. First, the companies that bid on reconnaissance aircraft all believed that the Air force was effectively the sole outlet for such pitches.[60] Interestingly, however, a relevant secret organization did exist. In July 1954, President Eisenhower tapped the President of MIT, James Killian, to head a group of scientific experts called the Technology Capabilities Panel (TCP) (Richelson, 2002, 11). Its existence was not widely known: "As with other secret panels formed by chief executives to deal with intelligence matters, Congressional input was missing from the TCP deliberations and few Congressmen knew it even existed, although many of its decisions had an immense impact on the nation's military and intelligence preparedness (5)."[61]

Project Three, one of three entities comprising the TCP, was a small group broadly focused on intelligence capabilities. It was not specifically tasked with developing proposals for overhead reconnaissance aircraft. Thus, the small and secretive Project Three members were not soliciting

---

[60]Although the CIA had developed several branches to deal with scientific intelligence and research and development in the early- to mid-1950s, they did not have much experience at that time with technical collection systems. See Fischer 2001 [OTS history].

[61]https://www.nro.gov/Portals/65/documents/history/csnr/programs/docs/prog-hist-04.pdf.

bids for such aircraft, and nobody expected that they would. However, the extreme secrecy that surrounded Project Three meant that they could develop research ideas in small teams that outsiders would not know about. Thus, unlike the Air Force, they exhibit the internal secrecy that our theory requires for secret innovation.

Based on this context, it is reasonable to assume that the Wright Air Development Command and any other relevant Air Force-related entity would hear all bids pertaining to overhead reconnaissance and had first right of refusal. Moreover, any project they did fund would at least be scrutinized by the broader Air Force leadership and possibly Congress. They would have also likely believed that anything they rejected would not be funded. However, as just noted, Project Three was quietly lurking in the background and ready to pick up rejected proposals if they so chose. This allows us to evaluate our counterfactual because we can observe: (1) what the open institution actually chose to accept and reject and; (2) given what the open institution rejected, what the secret institution chose to accept and reject.

C.2   Who funds what and why

The Air Force opted to pursue two proposals, the modified version of the B-57 from Martin which was viewed as a short-term solution and the Bell X-16 which promised better results in the medium-term. Bell was contracted to produce 28 such aircraft. At the same time, the Air Force rejected the Fairchild and Lockheed proposals. The Fairchild proposal was relegated to the dustbin of history. The Lockheed proposal was not. Lockheed took their proposal to various parts of the Air Force—including the Wright Air Development Command as well as Strategic Air Command and the Office of Development Planning—all of whom rejected it (Pedlow and Welzenbach, 1992, 11-12). Along the way, Project Three members learned of the Lockheed proposal and were immediately interested in it (Pedlow and Welzenbach, 1992, 31). As we will detail more in a moment, they undertook intense secretive research into CL-282's viability and verified that it would work. This project was later handed to the CIA as the U-2 project.

We predict that open organizations facilitate innovation when the benefits are clear (e_0 is positive), and there is not much disagreement about the likely effects (σ_0 is low); they will reject ideas that are radically new because they know little about them. Even though new ideas could have benefits, they could also cause damage. Open institutions are unlikely to take on projects like this even in the research phase (e_0 is near 0). Of these ideas, we predict that secretive institutions will pick them up as research projects if the potential outcomes vary widely (σ_0 is high). That is, there is a risk of catastrophic damage towards mission objects and enormous benefits that extend beyond what the other proposals could accomplish.[62]

This is precisely what we find. The Air force funded two safer projects that incrementally advanced the state of overflight. The modified B-57 is an obvious example. The goal was to "improv[e] the already exceptional high-altitude performance of the B-57 Canberra" (Pedlow and Welzenbach, 1992, 9). It "featured lengthened wings, accommodations for cameras and sensors, and uprated twin engines."[63] The Bell X-16 was slightly more advanced than the B-57. The modifications made to reduce weight and reach higher altitudes were far less radical than the CL-282.[64]

---

[62]That is, there is uncertainty about whether the innovation will move the U.S. closer to or further from its policy objectives.

[63]https://www.nasa.gov/sites/default/files/atoms/files/unlimited-horizons.pdf [page 1].

[64]https://www.nasa.gov/sites/default/files/atoms/files/unlimited-horizons.pdf [4-5].

7

The U-2 was radical by design. Senior Lockheed designers prioritized "nonstandard" elements, including "the elimination of landing gear, the disregard for military specifications, and the use of very low load factors" (Pedlow and Welzenbach, 1992, 10). Several elements of what was eventually dubbed the CL-282, and would later become the U-2, "were adapted from gliders. Thus, the wings and tail were detachable. Instead of conventional landing gear," Kelly Johnson, the lead developer, "proposed using two skis and a reinforced belly rib for landing—a common sailplane technique—and a jettisonable wheeled dolly for takeoff." As a declassified history of the U-2 puts it, "Essentially, Kelly Johnson had designed a jet-propelled glider" (Pedlow and Welzenbach, 1992, 12).

Part of Seaberg and the Wright Air Development Command's rationale for rejecting the CL-282 proposal speaks to their uncertainty about whether it would work. Seaberg pointed to its use "of the unproven General Electric J73 engine. The engineers at Wright Field considered the Pratt and Whitney J57 to be the most powerful engine available." All three of the other proposals they received from small manufacturers relied on the latter. Moreover, Seaberg and colleagues viewed "[t]he absence of conventional landing gear" on the CL-282 as a "shortcoming." Because the other proposals, including the most promising—the Bell—had "normal landing gear," they were considered "more conventional aircraft" (Pedlow and Welzenbach, 1992, 12-15).

Other Air Force commands also registered dismay at the novel features of CL-282. General Curtis LeMay, the head of Strategic Air Command, apparently "stood up halfway through the briefing, took his cigar out of his mouth, and told briefers, that if he wanted high-altitude photographs, he would put cameras in his B-36 bombers and added that he was not interested in a plane that had no wheels or guns." He called the meeting "a waste of his time" (Pedlow and Welzenbach, 1992, 12).[65]

According to the declassified history of the U-2, another driving factor in the Air Force's rejection of the CL-282 had to do with their "preference for multi-engine aircraft." This was based on familiarity and their experience with multi-engine aircraft during World War II and likely explains why they also opted for the Bell and Martin designed but rejected the Fairchild bid, which relied on a single engine. Moreover, "aerial photography experts" at the time "emphasized focal length as the primary factor in reconnaissance photography and, therefore, preferred large aircraft capable of accommodating long focal-length cameras" (Pedlow and Welzenbach, 1992, 13)

As the foregoing makes clear, the CL-282's novel design meant that many in the Air Force were skeptical about its chances of success. In terms of the model's parameters, the balance of Air Force staff thought the overall impact of the project would cause no benefit (or harm) for surveilling the Soviet Union and ultimately ensuring peace. However, some raised concerns which implied that it could have catastrophic effects: "there was the feeling shared by many Air Force officers that two engines are always better than one because, if one fails, there is a spare to get the aircraft back to base... Furthermore, a high-altitude reconnaissance aircraft deep in enemy territory would have little chance of returning if one of the engines failed, forcing the aircraft to descend" (Pedlow and Welzenbach, 1992, 13). In other words, there was concern that a single-engine plane that was missing key parts could crash inside the Soviet Union and conceivably spark a conflict.

To be sure, not everyone in the Air Force shared the view that the Bell and Martin proposals were superior to the CL-282. Trevor Gardner, Special Assistant for Research and Development, and some other officials thought it had potential. They believed "it gave promise of flying higher than

---

[65]LeMay's reaction illustrates one way that military culture imposes costs on innovators. As we argued, this makes innovation difficult in open institutions.

the other designs and because at maximum altitude its smaller radar cross-section might make it invisible to existing Soviet radars" (Pedlow and Welzenbach, 1992, 15). Thus, if it worked, its value would be larger than the other projects.

Taken together, these divergent views support the notion that there was deep uncertainty about what CL-282 would accomplish. While some believed it was unlikely to work and therefore have no effect, others thought it could have either very negative or very positive (i.e. more positive than the other designs) effects. If the Air Force had been the only organization that could have considered the overflight proposals, one of the most important innovations of the twentieth century may never have seen the light of day (Pocock, 2000, 14).

Project Three members were themselves sensitive to the risks associated overflight over the Soviet Union.[66] But despite these risks, they pursued the project because of the enormous potential upside if the project was successful. "By the end of October [1954], the Project Three meetings had covered every aspect of the Lockheed design. The CL-282 was to be more than an airplane with a camera, it was to be an integrated intelligence-collection system that the Project Three members were confident could find and photograph the Soviet Union's Bison bomber fleet and, thus, resolve the growing 'bomber gap' controversy." They were also taken with the prospect that the proposal could be "the platform for a whole new generation of aerial cameras" (Pedlow and Welzenbach, 1992, 31).

Their approach to research supports our theory in two additional ways. First, they operated in secret. Land and his team "began developing it into a complete reconnaissance system," meeting in small-group settings with usually less than 10 people present. Second, they did not instantly recommend production of U-2 planes. Rather, they exploited secrecy to determine if the project was viable. Once they realized it was, they revealed what they had been doing to the CIA Director and to President Eisenhower who was extremely receptive. He "approv[ed] the development of the system, but ... stipulat[ed] that it should be handled in an unconventional way so that it would not become entangled in the bureaucracy of the Defense Department or troubled by rivalries among the services" (Pedlow and Welzenbach, 1992, 33).[67]

Interestingly, the project also helped the TCP realize that secret organizations like the CIA were well-suited to the task of overseeing radical innovations of this kind. As the TCP argued to CIA Director Allen Dulles in a memo, "this seems to us the kind of action and technique that is right for the contemporary version of the CIA; a modern and scientific way for an Agency that is always supposed to be looking, to do its looking. Quite strongly, we feel that you must always assert your first right to pioneer in scientific techniques for collecting intelligence... This present opportunity for aerial photography seems to us a fine place to start."[68]

D   National Security and Innovation Literature

Since our theoretical framework is closest to principal-agent theories of organizational innovation, we focus our review on that literature. We also review works in international relations and bureaucratic politics that help us justify changes in our assumptions. However, our paper has broad substantive

---

[66]See https://history.state.gov/historicaldocuments/frus1950-55Intel/d194.

[67]Interestingly, the Air Force eventually comes around to accepting the proposal but does not actually abandon their X-16 program until the U-2 was operational.

[68]https://nsarchive2.gwu.edu/NSAEBB/NSAEBB74/U2-03.pdf.

interest for scholars of innovation and security broadly defined. Here we review four different strands of this literature, explain how we connect and contribute to them:

1. Bureaucracy and barriers to and opportunities for military innovation;

2. Adaptation and military innovation;

3. Conflict processes and innovation, which can examine autocratic repression or terrorism and innovation;

4. The strategic implications of new technology.

Many of the concepts we describe intersect with these literatures. But we frequently arrive at surprising conclusions for all of them. In what follows, we explain how our theory intersects with these important literatures and clarify differences.

D.1   Barriers and opportunities for military innovation

A large literature in security and strategic studies examines military innovation. Many of these analyses begin with the premise that, despite the importance of innovation to national security, military innovation is rarer than we might expect it to be. Why? The answer, in brief, is that innovators face costs of different kinds. One common impediment is that militaries are "hierarchical, inflexible, and rigid" (Jungdahl and Macdonald, 2015, 467). As Grissom (2006, 919) argues in his review of this literature, most scholars argue that "military organizations are intrinsically inflexible, prone to stagnation, and fearful of change." What this means in practice is that individuals are often reluctant to suggest new ideas for professional or cultural reasons, and new ideas that do get proposed can often get shut down.

Despite these barriers, militaries sometimes innovate. Thus, another key task of this literature is to answer the following question: what explains how militaries can overcome bureaucratic inertia or military culture to innovate? Some argue that military organizations may innovate when they face external pressures from the outside, usually from civilians (Posen, 1984). Another is when senior members of the military re-conceptualize their tasks and create career paths for new officers that incentivize the embrace of this new way of thinking (Rosen, 1988). A third set of explanations focuses on cultural differences (Adamsky, 2010; Farrell and Terriff, 2002) According to one study, a "receptive culture" can facilitate new thinking and vice versa.[69] A fourth argues that innovation requires special incubators, which are "informal subunits established outside the hierarchy" where individuals can collaborate, try out ideas, and push the envelope. There are others (Grissom, 2006).

While each of these pathways are distinct in important ways, they all share a common strategic logic. First, individuals inside the military face barriers (i.e. costs) to innovation. Therefore, they either do not voice their ideas, or are unable to push their ideas through the military bureaucracy. This explains why innovation does not happen often. Second, opportunities for innovation arise when military leaders, or outsiders with power create incentives (i.e. lower the costs associated with pursuing innovation). Things like new pathways to promotion, visionary civilians that intervene

---

[69]Price (2014). Lee (2019) has shown, for example, that the Air Force's cultural preference for manned systems led it to reject innovations in drone technology for longer than would otherwise be the case if one were using a strictly rationale model.

to support and defend new ways of doing business, and incubators where individuals can test ideas outside the formal process are a way for would-be innovators to safely conceive of ideas, develop them, and potentially implement them without incurring significant costs. Without these cost-lowering mechanisms, the argument goes, innovation does not happen.

Our theory accounts for these conditions in the costs and benefits parameters. The logic of our model under a specific set of parameters is consistent with the logic of these arguments. We find that researchers will not *openly* pursue innovation even when the policy implications are important (the expectation of pi is positive) if the organization imposes large personal costs on the agents.

The critical difference between our theory and this literature is what happens when the costs and benefits are high. Scholars of military innovation typically argue that if the costs of pursuing research are high then the innovators simply do not pursue their ideas. As noted, their logics for military innovation largely follow a similar process: some kind of organizational change transpires that lowers the costs associated with agents openly pursuing innovation; the researcher realizes that the organization is accommodating of new ideas; the researcher then raises their ideas with their manager so that they can openly pursue them. In our theory, national security researchers sometimes face another option: secret innovation. Rather than taking their idea to their manager, or sharing it broadly with others in their organization, a small team of researchers can pursue an idea in secret. This gives the researcher autonomy to pursue their idea and demonstrate its plausibility. It also allows different agents to distribute the high institutional costs associated with pursuing new ideas.

In this way, our theory illuminates that existing studies emphasize open, national security innovation in the way that we define openness.[70] As written in the manuscript, open research refers to a setting where individuals broadly share their ideas with their managers, people with budgetary oversight, and many others across their organization and possibly outside their organization.[71] What is more the costs that these scholars describe usually stem from openness. Consider that bureaucratic inertia, or cultural barriers only prevent pilot testing if ideas are shared openly. If a small team of researchers does not ask permission, they do not face bureaucratic inertia.

There are several other ways in which our theory differs from, but complements, broader literature on military innovation which includes both doctrinal innovations as well as technological and tactical innovations (Beard, 1976; Jungdahl and Macdonald, 2015; Sapolsky, 1972) First, most of these accounts emphasize innovation that occurs through a top-down process. Our focus entails a heavy bottom-up component (Griffin, 2017, 214).[72] Second, much of this scholarship on military innovation has a bias towards *successful* innovations.[73] By focusing on the process or pursuit of

---

[70]Scholars such as Kurth Cronin (2020, 23-28) discuss "closed innovation," defined as "state organizations creat[ing] and control[ing] high-end military technologies" such as nuclear weapons. Even in this case, though, while innovation may be hidden from the *outside* world it is still open internally within the government.

[71]Although they do not usually describe it this way, the existing security studies literature usually focuses on open innovation under this definition. Perhaps the clearest example of this is innovations in doctrine, a common focus of this literature. When doctrinal innovation happens, it is usually carried out in broad view of many parts of the military. It requires many services and branches to work together. Even during periods of conceptualization, new doctrine requires combat experts to interface with logistics, strategic intelligence, manpower and budget experts, defense contractors, and more. Moreover, since new doctrine requires new field manuals, soldiers tend to find out important details of doctrine as it is being developed.

[72]For exceptions, see Jungdahl and Macdonald (2015); Kollars (2014).

[73]This is evidenced by the way many scholars define innovation, which often requires things like improvements in military effectiveness. See Grissom (2006, 907). As Posen (1984, 29) notes, however, "Neither innovation nor stagnation ... should be valued a priori.

11

innovation, our study allows for the prospect that many of these ideas, particularly those pursued in secret organizations, will fail.

D.2 Adaptation and military innovation

A second literature examines diffusion and adaptation. This is similar because it examines military innovation. However, they focus on how existing military technologies diffuse cross-nationally. Horowitz (2010a, 3), for example, develops the "adoption-capacity theory" to explain "why some military innovations spread and influence international politics while others do not, or do so in very different ways." In a somewhat similar vein, Gilli and Gilli (2019, 141) examine the logic of imitation, asking whether America's rivals can "easily imitate its most advanced weapon systems and thus erode its military-technological superiority."

The aspect of these studies that is most similar to ours examines different ways that states adopt the same technology. This could be thought of as tactical innovations. However, these tactical innovations are typically described as open, and the primary barriers is in adopting an existing technology and not in finding new ways to use it.

D.3 Innovation among autocrats and terrorist groups

Our framework also differs from a newer literature on innovation among terrorist organizations and autocratic regimes. Regarding terrorist groups, innovation is often driven by the need to evade a target's defenses, amplify lethality, and shape public opinion (Horowitz, Perkoski and Potter, 2018). The precise characteristics of terrorist organizations, their leaders, and their broader environment, however, shape whether they are successful.[74] One of the key differences between these studies and our own is that terrorist organizations as a whole are insensitive to the costs of innovation whereas the individuals in our model are political actors and researchers with an entirely different incentive structure.[75]

Finally, there is an emerging literature that examines innovation and autocratic regimes. A key focus of these works is how dictators can exploit technological innovations to their advantage. This includes the use of the Internet and other technologies for the purposes of repression and surveillance (Dragu and Lupu, 2021; Gohdes, 2020). In these studies, autocratic leaders are exploiting existing technologies that may have been developed with an entirely separate purpose in mind for their own ends, including regime survival and population control. Like terrorist organizations, they are also insensitive to costs. As noted, our focus is on the sources of innovation in a situation where there are political actors who can distribute costs to subordinates.

D.4 Strategic implications of emerging technology

A growing litreature emphasizes the strategic implications of emerging technology (see Sechser et al., 2019, for review). We partly use this litearture to justify our claim that the benefits of innovation

---

[74]See Moghadam (2013); Perkoski (2019).

[75]To be sure, terrorists may be sensitive to how the public will *perceive* an innovation such as suicide bombing but are themselves by and large insensitive or at least willing to incur enormous costs given the nature of asymmetric conflict.

(i.e. whether innovation moves you closer or further from your policy goals) is uncertain. This literature is more about what states do with innovations once they have them. It is less about why states decide to pursue them in the first place (Garfinkel and Dafoe, 2019; Horowitz, 2016; Zhang et al., 2021).

E   Principal-agent literature

Our substantive focus is foreign policy and international relations. However, as we discuss in the manuscript, the structure of our theory is closest to principal-agent theories of organizational innovation in the private sector (Lai et al., 2009). These theories emphasize aspects of PA problems not commonly studied by international relations scholars. In what follows, we explain how our theory fits within the PA framework. We then clarify important differences with three applications of PA theory in IR.

E.1   What makes our theory a principal agent theory?

PA theory is very broad (Eisenhardt, 1989). There are many types of principal-agent problems that scholars study including moral hazard, agency loss, adverse selection, credible communication, and unjust reprisals (Stiglitz, 1989; Hart and Holmström, 1987). While each problem is different, they are united by a few common elements. In this section, we describe the elements of a PA theory and how our theory includes these elements.

A basic principal-agent dynamic (or contract theory) involves at least one agent and at least one unified principal that have asymmetric preferences and in which the agent is given a choice to impact the principal's welfare (Miller, 2005). Our basic institution models these elements. We study a researcher and manager who vary in their cost functions. As a result of these cost functions, situations arise where the researcher wants to pursue research and development but the manager does not. We make one assumption that is common in models of innovation: the effects of pursuing a policy follow from imperfect information and are not known to either player. This assumption is not common in PA models of policy-making (e.g. Downs and Rocke, 1994). The reason is that policy-makers (i.e. the agent) knows whether their choice will benefit the principal with a large degree of confidence (i.e the public); at least ex-post.

Beyond this difference, we make a novel assumption in the basic model that departs from PA models of innovation: the researchers can exploit secrecy to distribute costs. This creates a dynamic in which the researcher can incur costs to pursue outcomes that the manager would veto. We study the impact of this additional assumption under complete information because it generates a novel tension not typically appreciated in PA models.

Principal-agent theories introduce problems through asymmetric information, and a principal's initiative (Miller, 2005). The specific type of principal-agent problem varies depending on how scholars introduce private information (Hart and Holmström, 1987). We model two variants of a principal-agent problem in extensions 3.3.1 and 3.3.2. The first represents a monitoring problem, the second represents a credible advice problem. Past scholars examine how variation in the costs of monitoring, agent selection, or punishments can elicit agency compliance and the credible revelation of information. However, we find that secrecy paradoxically alleviates many of the common

problems of asymmetric preferences and information. It also creates new incentives for the manager to extract value from the researcher's compliance.

E.2   How is this different from PA models in international relations and foreign policy studies?

Here we describe three literatures that examine principal agent problems in international relations: hierarchy, security force assistance, and gambling for resurrection.

We start with a joint-discussion of hierarchy (Hawkins et al., 2006; Nielson and Tierney, 2003; Lake, 2001) and security force assistance (Biddle et al., 2018; Ladwig, 2016). Of course, these empirical domains are very different from each other. Further, each domain includes many different studies that tackle different aspects of the PA problem. However, they are all united by the fact that they assume the principal and agent come from different states and therefore have dramatically different preferences. Scholars of security force assistance assume that the principal is either US military advisers or the entire US military and the agent is the military of another state (e.g. the Afghan army).

We do not focus on a situation like this. Consistent with organizational models of principal-agent theory and innovation, we examine individual employees (or small groups of individuals) who work at a single organization (or a handful of closely connected agencies that share a common mission within the executive branch of a single country; like the CIA and NRO). To match this domain, we assume that the researcher and manager both share an interest in advancing the organization's overall goals (both researcher and manager's utility is increasing in π). However, their preferences over research and development still vary because the personal and professional incentives of managers and researchers vary (c, k can vary).

Our assumptions are appropriate for the setting we study. The goals that national security agencies pursue are things like defeating the Soviet Union in the Cold War, or winning the Second World War. In general, we believe that managers and researchers employed in the national security community benefit to the extent that they succeed in these goals and lose to the extent they fail in them. This is partly due to the extensive security clearance process and constant monitoring that national security employees are subject to. It also relates to professional incentives once in these communities. Finally, evidence suggests that public sector employees, and especially national security employees, tend to have a strong public service motivation. However, individual agents may disagree about the best way to achieve these goals, face incentives to buck-pass, or have parochial incentives that cause them to weight the costs and benefits differently.

Studies of gambling for resurrection are closer to us because they examine a leader and the public of the same country. Most notably, Downs and Rocke (1994) theorize about the president as the agent who makes the choice to fight a war (or not). The president holds asymmetric information over whether war serves the public interest. They model the public as the principal who can re-elect the president. This model is closer to ours than the hierarchy and security force assistance literatures in that the public and the president both share a preference for avoiding bad foreign policy outcomes.

But there are several differences. First, the president has a unique incentive for re-election that can conflict with the public's. As discussed, these preferences are not appropriate in our theory

(although our theory is robust if we model preference variation like this). Second, the president has private information about the quality of the choice to fight, and his own quality. This is not appropriate in our model for two reasons. The first reason is, unlike the American public, the manager has a security clearance and access to a wide cadre of classified researchers who can review the existing data. The second reason is that the researcher is very uncertain before they engage in pilot research precisely because they have not worked on a problem like this. Third, the public directly punishes the president through an electoral mechanism. This is not appropriate in our theory for two reasons. One reason is that the manager is complicit through don't-ask-don't tell, and therefore does not do the punishing. Another is that punishment does not take the form of replacing a researcher with a different one (as in the electoral context).

References

Adamsky, Dima. 2010. *The Culture of Military Innovation: The Impact of Cultural Factors on the Revolution in Military Affairs in Russia, the US, and Israel.* Stanford: Stanford University Press.

Beard, Edumnd. 1976. *Developing the ICBM: A Study in Bureaucratic Politics.* New York: Columbia University Press.

Biddle, Stephen, Julia Macdonald and Ryan Baker. 2018. "Small footprint, small payoff: The military effectiveness of security force assistance." *Journal of Strategic Studies* 41(1-2):89–142.

Downs, George W. and David M. Rocke. 1994. "Conflict, Agency, and Gambling for Resurrection: The Principal-Agent Problem Goes to War." *American Journal of Political Science* 38(2):362.

Dragu, Tiberiu and Yonatan Lupu. 2021. "Digital Authoritarianism and the Future of Human Rights." *International Organization* 75(4):991–1017.

Eisenhardt, Kathleen M. 1989. "Agency Theory: An Assessment and Review." *The Academy of Management Review* 14(1):57.

Farrell, Theo G. and Terry Terriff. 2002. *The sources of military change: Culture, politics, technology.* Boulder, CO: Lynne Rienner.

Garfinkel, Ben and Allan Dafoe. 2019. "How does the offense-defense balance scale?" *Journal of Strategic Studies* 42(6):736–763.

Gilli, Andrea and Mauro Gilli. 2019. "Why China Has Not Caught Up Yet: Military-Technological Superiority and the Limits of Imitation, Reverse Engineering, and Cyber Espionage." *International Security* 43(3):141–189.

Gohdes, Anita R. 2020. "Repression technology: Internet accessibility and state violence." *American Journal of Political Science* 64(3):488–503.

Griffin, Stuart. 2017. "Military Innovation Studies: Multidisciplinary or Lacking Discipline?" *Journal of Strategic Studies* 40(1-2):196–224.

Grissom, Adam. 2006. "The future of military innovation studies." *Journal of strategic studies* 29(5):905–934.

Hart, Oliver and Bengt Holmström. 1987. The theory of contracts. In *Advances in Economic Theory.* Cambridge University Press pp. 71–156.

Hawkins, D G, D A Lake, D L Nielson and M J Tierney. 2006. *Delegation and Agency in International Organizations.* Political Economy of Institutions and Decisions Cambridge University Press.

Horowitz, Michael. 2010. *The diffusion of military power : causes and consequences for international politics.* Princeton University Press.

Horowitz, Michael C. 2016. "Public Opinion and the Politics of the Killer Robots Debate." *Research & Politics* 3(1):1–8.

Horowitz, Michael C., Evan Perkoski and Philip B.K. Potter. 2018. "Tactical Diversity in Militant Violence." *International Organization* 72(1):1–35.

Jungdahl, Adam M and Julia M Macdonald. 2015. "Innovation inhibitors in war: Overcoming obstacles in the pursuit of military effectiveness." *Journal of Strategic Studies* 38(4):467–499.

Kollars, Nina. 2014. "Military innovation's dialectic: Gun trucks and rapid acquisition." *Security Studies* 23(4):787–813.

Kurth Cronin, Audrey. 2020. *Power to the People: How Open Technological Innovation is Arming Tomorrow's Terrorists.* New York: Oxford University Press.

Ladwig, Walter C. 2016. "Influencing Clients in Counterinsurgency: U.S. Involvement in El Salvador's Civil War, 1979–92." *International Security* 41(1):99–146.

Lai, Edwin L.-C., Raymond Riezman and Ping Wang. 2009. "Outsourcing of innovation." *Economic Theory* 38(3):485–515.

Lake, David a. 2001. "Beyond Anarchy: The Importance of Security Institutions." *International Security* 26(1):129–160.

Lee, Caitlin. 2019. "The Role of Culture in Military Innovation Studies: Lessons Learned from the US Air Force's Adoptino of the Predator Drone, 1993-1997." *Journal of Strategies Studies* pp. 1–35.

Miller, Gary J. 2005. "THE POLITICAL EVOLUTION OF PRINCIPAL-AGENT MODELS." *Annual Review of Political Science* 8(1):203–225.

Moghadam, Assaf. 2013. "How al Qaeda innovates." *Security Studies* 22(3):466–497.

Nielson, Daniel L. and Michael J. Tierney. 2003. "Delegation to International Organizations: Agency Theory and World Bank Environmental Reform." *International Organization* 57:241–276.

Perkoski, Evan. 2019. *Terrorist technological innovation.* Oxford: Oxford University Press.

Posen, Barry. 1984. *The sources of military doctrine: France, Britain, and Germany between the world wars.* Ithaca: Cornell University Press.

Price, John F. 2014. "US Military Innovation: Fostering Creativity in a Culture of Compliance." *Air & Space Power Journal* 43(Sep.-Oct.):128–134.

Rosen, Stephen Peter. 1988. "New ways of war: understanding military innovation." *International security* 13(1):134–168.

Sapolsky, Harvey M. 1972. *Polaris System Development: Bureaucratic and Programmatic Success in Government.* Cambridge, MA: Harvard University Press.

Sechser, Todd S., Neil Narang and Caitlin Talmadge. 2019. "Emerging technologies and strategic stability in peacetime, crisis, and war." *Journal of Strategic Studies* 42(6):727–735.

Stiglitz, Joseph E. 1989. Principal and Agent. In *Allocation, Information and Markets.* London: Palgrave Macmillan UK pp. 241–253.

Zhang, Baobao, Markus Anderljung, Lauren Kahn, Noemi Dreksler, Michael C. Horowitz and Allan Dafoe. 2021. "Ethics and Governance of Artificial Intelligence: Evidence from a Survey of Machine Learning Researchers." *Journal of Artificial Intelligence Research* 71:591–666–591–666.