

---

# Secret Innovation

Michael F. Joseph<sup>a\*</sup>  and Michael Poznansky<sup>b</sup> 

<sup>a</sup>Department of Political Science, University of California, San Diego, USA

<sup>b</sup>Strategic and Operational Research Department, US Naval War College, Newport, RI, USA

\*Corresponding author. Email: [mjoseph@ucsd.edu](mailto:mjoseph@ucsd.edu)

---

**Abstract** Conventional wisdom holds that open, collaborative, and transparent organizations are innovative. But some of the most radical innovations—satellites, lithium-iodine batteries, the internet—were conceived by small, secretive teams in national security agencies. Are these organizations more innovative because of their secrecy, or in spite of it? We study a principal–agent model of public-sector innovation. We give research teams a secret option and a public option during the initial testing and prototyping phase. Secrecy helps advance high-risk, high-reward projects through the early phase via a cost-passing mechanism. In open institutions, managers will not approve pilot research into high-risk, high-reward ideas for fear of 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 with two cases from the early Cold War: the CIA’s attempt to master mind control (MKULTRA) and the origins of the reconnaissance satellite (CORONA). We contribute to the political application of principal–agent theory and studies of national security innovation, emerging technologies, democratic oversight, the Sino–American technology debate, and great power competition.

---

Nations that want prosperity and security must innovate.<sup>1</sup> Scholars across many disciplines study why specific organizations are innovative. There is broad agreement that openness, defined loosely as organizations that encourage employees to internally share ideas with colleagues, spurs innovation.<sup>2</sup> Internally open organizations foster competition and collaboration between otherwise siloed divisions,<sup>3</sup> diffuse ideas,<sup>4</sup> and encourage a free flow of information.<sup>5</sup>

1. Horowitz and Pindyck 2023; Taylor 2016.
2. Laursen and Foss 2003; West and Anderson 1996.
3. Aghion et al. 2005; Macdonald 2015.
4. Boushey 2016.
5. Zoghi, Mohr, and Meyer 2010.

*International Organization* 78, Fall 2024, pp. 766–99

© The Author(s), 2024. Published by Cambridge University Press on behalf of The IO Foundation. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

doi:10.1017/S0020818324000250

One set of institutions consistently buck this trend: secretive intelligence and national security organizations. Agencies like the US's Central Intelligence Agency (CIA) and the Defense Advanced Research Projects Agency (DARPA), or the UK's MI6, often discourage internal sharing, but have consistently produced radical innovations. These include the satellite,<sup>6</sup> autonomous robots,<sup>7</sup> and lithium-iodine batteries.<sup>8</sup> Failed projects—from turning cats into listening devices, to psychic spying—also speak to their vision.<sup>9</sup>

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.<sup>10</sup> We adapt the payoffs to reflect the actors' sensitivities to political costs and benefits.<sup>11</sup> We allow researchers secrecy during the conceptualization of innovation. We then contrast mechanisms for innovation in open versus secret public-sector institutions.<sup>12</sup>

Secrecy allows actors to distribute the political costs of authorizing each phase of politically sensitive research. In open institutions, lower-level researchers cannot pursue pilot programs to determine whether a concept is viable without their manager knowing. When a manager learns of a novel but controversial idea, they will not even approve pilot research because they do not want to be responsible. Secrecy early in the innovation process gives an enterprising researcher cover to collect evidence (at a larger personal cost) that the novel idea is viable. If it shows promise, the researcher seeks manager approval.

Our mechanism generates two surprising results for international relations principal–agent theory.<sup>13</sup> First, the researcher turns to secrecy even if their preferences are perfectly aligned with the manager's. Second, the manager does not monitor the researcher even if monitoring is costless and perfect *and* the manager knows that the researcher is exploiting secrecy only to do something the manager would not allow them 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 that if they monitor the researcher they will discover something unsavory and shut the research down. However, if they remain ignorant, they can incur a small share of the costs associated with highly controversial pilot research and still benefit when it shows promise.

We find that, on average, secretive national security institutions should produce different kinds of innovations from open institutions. All political organizations pursue ideas that in expectation serve the national interest and involve

6. Oder, Fitzpatrick, and Worthman 1988.

7. Jacobsen 2015.

8. Richelson 2002, 257.

9. Houghton 2019; Richelson 2002.

10. Kopel and Riegler 2008; Lai, Riezman, and Wang 2009.

11. Joseph, Poznansky, and Spaniel 2022.

12. Cain 2015.

13. Di Lonardo, Sun, and Tyson 2020; Downs and Rocke 1994; Hawkins et al. 2006; Malis 2021, 2024.

uncontroversial research practices. However, only secret organizations can pursue *initial concepts* that involve large risks and rewards. Before pilot research is carried out, these ideas are too controversial or novel 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 several cases: attempts to master mind control (MKULTRA) and innovations to facilitate overhead reconnaissance (via the CORONA satellite and, in the online supplement, the U-2 spy plane). We chose these cases because of historical features that provide inferential leverage. They are also important for, but overlooked by, international relations scholars. MKULTRA's exposure in the 1970s affected intelligence reforms and legislative oversight for decades afterward. The case is especially important given that, according to secondary accounts, the CIA and DARPA recently looked into brainwashing.<sup>14</sup> Existing studies examine the effects of reconnaissance technologies on conflict dynamics,<sup>15</sup> but overlook what it took to develop them. We show that these research programs might have been shelved but for internal secrecy.

We contribute to several debates. First, we provide a logic for the national security origins of radical innovations. For international security scholars, this clarifies selection into technology shocks that can cause<sup>16</sup> and offset<sup>17</sup> conflict. It also helps explain the conditions under which state-led, highly productive innovations of interest to international political economy scholars occur.<sup>18</sup>

Second, we refine theories of military innovation.<sup>19</sup> Conventional wisdom holds that while "militaries have strong incentives to innovate in order to succeed in war," they are "slow to innovate" because of hierarchical structures and entrenched interests.<sup>20</sup> We connect agency-wide incentives to respond to international threats with individual-level incentives for innovation.<sup>21</sup> We highlight an unexplored bureaucratic feature: agencies that practice internal secrecy owing to fear of foreign rivals. Since the level of internal secrecy varies across organizations and time, we also explain unexplored variation in peacetime innovation. Moreover, we broaden this research to agencies beyond the military, and focus on technological rather than doctrinal innovation.

Third, we contribute to secrecy research in international relations. Most find that secrecy reduces welfare<sup>22</sup> because it creates uncertainty at the international

14. Jacobsen 2015, 109.

15. Carnegie and Carson 2018; Coe and Vaynman 2020; Early and Gartzke 2021; Vaynman 2022.

16. Debs and Monteiro 2014.

17. Coe and Vaynman 2020.

18. Drezner 2019; Farrell and Newman 2019; N.L. Miller 2022; Taylor 2016. Some estimate that satellites have contributed over a trillion dollars to the US economy. Microchips, lithium-ion batteries, and the green revolution hold similar implications.

19. Kuo 2022; Posen 1984; Rosen 1988.

20. Neads, Farrell, and Galbreath 2023.

21. Jungdahl and Macdonald 2015.

22. Carnegie 2021.

level<sup>23</sup> and facilitates domestic inefficiencies.<sup>24</sup> We expand the limited research on efficient secrecy by showing that it allows states to pursue welfare-enhancing projects, not only to avoid costs.<sup>25</sup> We also define internal and external secrecy and explain their connection.

## Concepts

Our theory is closest to principal–agent models that examine rationalist, organizational innovation.<sup>26</sup> We adapt this model to fit public-sector employees, and secrecy. Others examine principal–agent problems in political institutions,<sup>27</sup> but they typically focus on policymaking or electoral accountability, not innovation.<sup>28</sup> Some examine principal–agent problems unique to international relations, but they emphasize interactions between two states<sup>29</sup> or foreign militaries.<sup>30</sup> We focus on a handful of employees working within government agencies. We detail the differences in Appendix D (in the online supplement). Our theory shares a substantive focus with theories of national security innovation, but our approach is different. We review how we complement this literature in Appendix C. Here we further develop our two central concepts: innovation and internal secrecy.

### *Innovation*

Innovation is the process of taking a novel idea and converting it into a working device or policy.<sup>31</sup> Innovation occurs only after (1) a novel idea, (2) pilot testing to validate and improve that insight, and (3) the decision to develop a product and deploy it in the field.<sup>32</sup> The last step is critical. It is not enough to conceive an idea. Innovation requires that it is developed into a working product.<sup>33</sup>

Government agencies innovate to achieve their policy goals.<sup>34</sup> Consistent with others, we define an innovation's effects as whether the final product advances national goals such as military effectiveness, intelligence collection, and security and prosperity.<sup>35</sup>

23. Carnegie and Carson 2018; Joseph and Poznansky 2018; Wolford, Reiter, and Carrubba 2011.

24. Colaresi 2012; Goldfien and Joseph 2023; Goldfien, Joseph, and Krcmaric 2024.

25. Carson 2018; Joseph 2021; Kurizaki 2007; Poznansky 2020.

26. Kopel and Riegler 2008; Lai, Riezman, and Wang 2009.

27. G.J. Miller 2005.

28. Downs and Rocke 1994 is closest to the present study.

29. Hawkins et al. 2006.

30. Biddle, Macdonald, and Baker 2018.

31. Kollars 2017; Taylor 2016.

32. Horowitz and Pindyck 2023; King 1990.

33. West and Anderson 1996.

34. Taylor 2016.

35. Horowitz and Pindyck 2023. We bracket distributional concerns because our actors are national security professionals, who typically have stronger public service motivations than the average citizen. Houston 2000.

Many innovations have positive effects (that is, move the organization toward its goals). Others have no effect. And still others have negative effects (unintended consequences).<sup>36</sup> This could include escalating conflict, weakening defenses, or facilitating local rebellion.<sup>37</sup> While researchers hold expectations, they cannot be certain of the true outcome.

We distinguish between the effects of innovation and the costs of moving an idea through development phases.<sup>38</sup> Some development costs stem from the financial burden of research trials and prototype construction. But public-sector institutions are especially sensitive to political costs.<sup>39</sup> These can manifest at different stages of innovation. During pilot research, political costs can come from wasteful spending or from human subjects research without consent. During the deployment phase, they can follow from labor abuses during production or escalation with rivals.

Of course, not all research activates political costs.<sup>40</sup> But in many cases, national security employees do face costs, including from organizational cultures that perceive radical ideas as reckless.<sup>41</sup> Since they are dealing with public funds, risky spending (which is promoted in private firms) is often viewed as a violation of the federal code of conduct under the waste, fraud, and abuse standard. Many ideas also raise the risk of tragic accidents in which soldiers die or test satellites crash into foreign territory. When this happens, investigators scrutinize those who plan and approve these programs looking for mistakes. These anticipated costs can be large enough that researchers do not voice their ideas in the first place. This helps explain why militaries may fail to pursue novel ideas even though the problems are important and their budgets are large.

Later, these contextualizing details will help us interpret our theoretical findings. But in the end, our model is abstract. We assume only that different public-sector employees participate in the research process and derive benefits (positive or negative) depending on the effects of innovation. They also incur research and development costs as ideas go through the innovation process. The scope of these costs depends on how responsible they are for advancing an idea, and on their personal sensitivities.

### *Internal Secrecy*

While democracies promote scrutiny of government agencies, national security agencies enjoy a special status. Specifically, they are allowed to keep secrets owing to fears of foreign threats.<sup>42</sup> We refer to this phenomenon as external secrecy. But to

36. Joseph 2023; Kuo 2022; Sechser, Narang, and Talmadge 2019.

37. Horowitz 2020; Kuo 2021.

38. King 1990.

39. Private and public organizations both accrue political and financial costs. However, firms mainly consider the financial liabilities of both costs and effects. *Ibid.*

40. Our model accounts for this because we allow costs to be zero.

41. Grissom 2006; Lee 2023.

42. Colaresi 2014.

sustain external secrecy, they also often practice internal secrecy. That is, certain individuals and groups are exempt or discouraged from sharing information with colleagues, those who oversee them, or even their own superiors.<sup>43</sup>

Internal secrecy is necessary to sustain external secrecy for two reasons. First, foreign threats may infiltrate national security agencies. Foreign penetration can be stemmed by limiting who knows important facts. Second, whistleblowers and leakers have historically revealed large document corpuses without fully understanding the potential for national security harm. By restricting access, agencies limit what they release publicly.<sup>44</sup>

Some aspects of internal secrecy are institutionalized. In the United States, for example, program officers must alert contract officers to purchases, who then openly tender contracts. But “full and open competition need not be provided for when the disclosure of the agency’s needs would compromise the national security.”<sup>45</sup> Other aspects of internal secrecy follow from a culture of need-to-know. Because of the “sensitive nature of their work, intelligence organizations have been reluctant to engage in bidirectional dialogue with decision makers and the larger public.”<sup>46</sup>

Both monitoring and evaluation, and budget oversight, are mandatory for most public-sector agencies. But secretive intelligence agencies and certain parts of the military have access to unvouchered funds they can spend on research without explaining what it is for.<sup>47</sup> According to a senior Government Accountability Office official, “we have no access to certain CIA ‘unvouchered’ accounts and cannot compel our access to foreign intelligence and counterintelligence information ... We have not actively audited the CIA since the early 1960s.”<sup>48</sup> A former CIA historian notes that “scrutiny of the [intelligence] budget ranged between ‘cursory and nonexistent’.”<sup>49</sup>

One consequence of internal secrecy is that managers are partly forgiven for their ignorance when subordinates do things the manager did not expect.<sup>50</sup> 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 employees. But national security employees are expected to maintain secrecy to guard against leaks and counterintelligence threats. This helps excuse managers who do not intrusively monitor their staff to learn about questionable choices. During the

43. Delegation is an important part of secrecy in our theory. Others have studied delegation in private-sector innovation but find it only enhances innovation because information is liberally shared. Aghion et al. 2005; Jones, Kalmi, and Kauhanen 2006; West and Anderson 1996.

44. CIA 1960.

45. Subpart 6.3, Federal Acquisition Regulation, <<https://www.acquisition.gov/far/subpart-6.3>>.

46. Ivan, Chiru, and Arcos 2021, 505.

47. Johnson 2022, 168; see also Jacobsen 2015, 253.

48. Hinton 2001, 1.

49. Haines 1999, 85.

50. We parameterize the extent to which managers are forgiven as  $x$ .

Iran-Contra affair, for instance, President Reagan avoided some of the worst costs by claiming that subordinates engineered the scheme without his knowledge.<sup>51</sup>

We are interested in how internal secrecy impacts innovation in national security agencies. In short, secrecy is most salient during the early phases of innovation: periods where researchers develop prototypes or run laboratory tests and simulations without a manager or compliance officer knowing about it. Secondary accounts of DARPA program managers “start[ing], continu[ing], or stop[ping] research projects with little outside intervention” is a prime example.<sup>52</sup> As projects progress, even secretive agencies may exploit open research practices to refine ideas by sharing information broadly across the national security community. But in the absence of small teams pursuing initial testing in relative secrecy, many innovations may never make it that far.

To be clear, there are other parts of government with some internal secrecy. For example, in parliamentary democracies, cabinet documents are sealed for decades so that elected leaders can brainstorm policy innovations.<sup>53</sup> But this secrecy is confined to top-level policy discussion and does not cover the design and testing of products. Pilot studies and focus group research to support policies formulated by the cabinet are not privileged.

In practice, actors 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 in who knows devilish details and who does not. At one extreme, a handful of scientists know the controversial research activities but even their immediate superiors are unaware. At the other extreme, the executive is fully aware of the devilish details but legislators are not. In the middle, directors of intelligence agencies may know exactly what their subordinates are doing but not inform the executive.<sup>54</sup> 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 without explaining their practices outside the group and who share the costs of authorization if things go wrong, and *outsiders*, who can escape some costs by remaining ignorant about what subordinates are up to but cannot stop programs for long.

## 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 versus open organizations to explain the core mechanism driving secret innovation. Third, we use comparative statics to explore the innovations uniquely pursued in

51. Byrne 2014.

52. Jacobsen 2015, 6.

53. Cain 2015.

54. In other examples there are interagency 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.

secret organizations. Fourth, we introduce two distinct information, agency, and monitoring problems to flesh out the mechanism and connect the model to the principal–agent literature. Finally, we consider the rationale for allowing secrecy given that it can lead to perverse outcomes.

### Setup

We study an institution that employs two actors: a researcher (R, she) and a manager (D for decider, he). Figure 1 visualizes the game tree and payoffs. The dashed box is the subgame in which R exploits secrecy. In it, she can conduct pilot research without her manager knowing. Later, we contrast secret and open institutions. Open institutions remove the secret subgame but are otherwise identical.

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

Along the way to innovation, actors can authorize pilot research, which has two effects. First, it 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  $m \sim N(\pi, \sigma)$  tied to the true consequences of innovation ( $\pi$ ).<sup>55</sup>

Actors pay political costs for participating in a controversial research process. We assume players pay one cost,  $k_i$ ,  $i \in \{R, D\}$ , if the institution engages in pilot research.<sup>56</sup> They pay a second cost,  $c_i$ , if the project is deployed in 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 is  $1 + x$ . We distribute 1 unit of cost to the actor that chooses to take costly action (conduct research, authorize innovation), and a smaller portion  $x \in (0, 1)$  to the other actor who works at the same institution but did not directly take a costly action.<sup>57</sup>

### Analysis: Secret Innovation and the Cost-Passing Mechanism

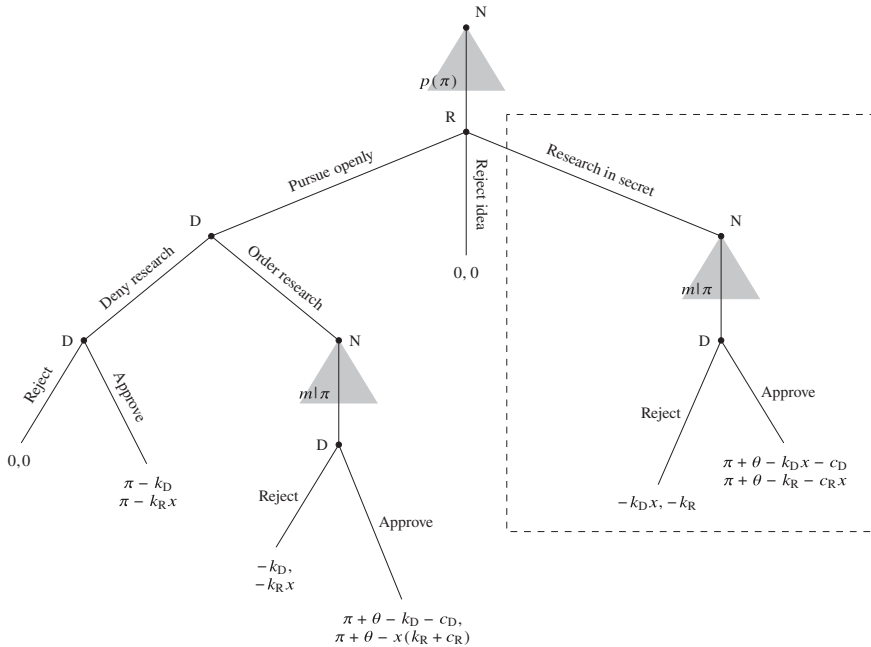
We solve for subgame perfect equilibria in the main model and extensions unless otherwise stated. We define secret innovation as follows.

55. Note  $\pi$  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.

56. A cost of zero implies that pilot research is not controversial.

57. As is standard,  $c$ ,  $k$  represent an actor's cumulative expectation of harm incurred at the moment a choice is made, based on the likelihood each possible punishment will be imposed and agents' sensitivities to each punishment.  $x$  allows authorization and knowledge of wrongdoing to moderate this expectation.





Parameter	Summarized interpretation
$\pi \in \mathbb{R}$	Effect of unleashing an innovation on the world
$p(\pi)$	Initial expectation about the innovation's effect
$c_i \geq 0$	Expected political costs actor $i$ incurs when innovation is approved
$k_i \geq 0$	Expected political costs actor $i$ incurs when pilot research is approved
$x \in [0, 1)$	What actor pays who participates in innovation phase that it did not authorize
$m() \pi$	What is learned about $\pi$ from pilot research
$\theta \geq 0$	How much pilot research improves an idea

Notes: Dashed rectangle represents the secret option. The open organization omits this subgame. Shaded triangles represent random variables. Nature does not reveal  $\pi$  to either player. Nature reveals  $m$  to both players.

FIGURE 1. Game tree and payoffs for baseline model

*Definition:* Suppose a fixed set of parameter values in the open institution where innovation cannot occur with probability on the path in any equilibrium. Then secret research facilitates innovation if innovation 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. There are two potential pathways. First, D can innovate absent research. Define  $e_0$  as the actors' prior expected value of  $\pi$ . Second, D can research and then decide to innovate if the research shows sufficient promise. We define two expectations at the moment

D must authorize research (or not). Define  $\lambda = pr(\mathbb{E}[\pi|m] > c_D - \theta)$ . Informally, this is D's pre-research belief that if research is conducted, he will observe a signal  $m$  that will lead to a posterior belief that the project is sufficiently likely to have benefits that outweigh the costs ( $\mathbb{E}[\pi|m] > c_D - \theta$ ). That is, it is D's pre-research belief that D will innovate after observing research. Define  $e_1 = \mathbb{E}[\mathbb{E}[\pi|m]|\mathbb{E}[\pi|m] > c_D - \theta]$ . Informally, this is D's pre-research expected value of  $\pi$ , given that D will observe an  $m$  sufficiently large that D is willing to approve research. Appendix A.1 gives more technical information on these expectations.

*Lemma 1: Neither research nor innovation can happen in the open institution if*

$$e_0 < k_D \tag{1}$$

and

$$\lambda < \frac{k_D}{e_1 + \theta - c_D} \tag{2}$$

*are satisfied. In every subgame perfect equilibrium player utilities are  $U^D = U^D = 0$ .*

See Appendix A.2. If condition 2 was violated, D would conduct research to determine whether the project is viable. But two factors drive D to reject a request for pilot research. First, research involves political costs ( $k$ ).<sup>58</sup> 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 1: Secrecy facilitates innovation if conditions 1, 2 and*

$$\frac{k_R}{e_1 + \theta - c_{RX}} < \lambda \tag{3}$$

*are satisfied. If they are, then in every subgame perfect equilibrium, R exploits secrecy to conduct pilot research, and 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.3. The result describes a condition where the researcher is willing to exploit secrecy to conduct research (condition 3 is satisfied), but her manager was unwilling to approve open research (condition 2 is satisfied). If research provides

58. Trivially, if research is costless or beneficial you always see open innovation.

evidence the project is viable ( $m$  suggests  $\pi$  is higher than originally thought), the manager will approve the project, leading to innovation.

Notice that we can achieve secret research even if the manager's and the researcher's cost parameters 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 exploit secrecy only 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 her secret pilot research shows promise ( $m$  is large), she can take the results to her manager for approval. Thus, the researcher is willing to assume the up-front cost and risk of research because she can convince her manager to bear the brunt of deployment costs.

### *Predictions About Ideas: Secrecy Drives Innovation When Initial Ideas Are High-Risk, High-Reward*

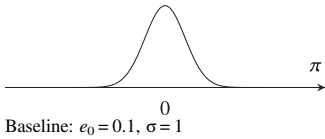
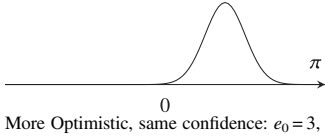
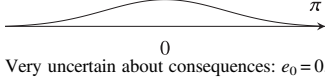
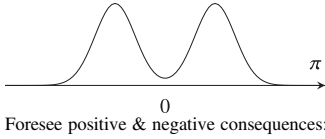
What are the kinds of initial ideas researchers need secrecy to pursue? Using a comparative static analysis, we expose two ideal-type pathways to secret innovation that are made possible because the manager and researcher weigh certain trade-offs differently. We provide technical support for these pathways in Appendix A.4. We visualize the results in Table 1. These pathways can interact. However, the basic trade-offs we identify are always present. Thus, it is valuable to consider them as distinct.

The first pathway appreciates the actor'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 other three panels represent different ways initial beliefs can vary.

First, they vary in their average expected effects,  $e_0$ . As  $e_0$  increases (row 2), it means the institution initially sees the idea as increasingly likely to yield a net benefit if it is developed and deployed in the field. The second way initial ideas vary is in the standard error of  $p()$ . We notate it  $\sigma$ . 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),  $\sigma$  represents disagreement about the potential consequences of innovation. The debate surrounding autonomous weapons is instructive. Proponents emphasize greater speed and stealth with fewer casualties. Critics point out that they might create greater instability and more crises.<sup>59</sup> Before these systems are deployed, it is hard to know whether they will provide benefit or cause harm.

59. Laird 2020; Zhang et al. 2021.

**TABLE 1.** *The innovation pathways for different initial ideas*

Initial predictions of effects, $p(\pi)$	Substantive description	$k_D$ low	$k_D$ high
 <p>Baseline: <math>e_0 = 0.1, \sigma = 1</math></p>	Project team largely agrees innovation won't impact state goals.	Scrap idea	Scrap idea
 <p>More Optimistic, same confidence: <math>e_0 = 3, \sigma = 1</math></p>	Project team largely agrees innovation will positively impact state goals.	Open research	Secret research <sup>2</sup>
 <p>Very uncertain about consequences: <math>e_0 = 0.1, \sigma = 3</math></p>	Project team is widely uncertain about project effects.	Secret research <sup>1</sup>	Scrap idea
 <p>Foresee positive &amp; negative consequences: <math>e_0 = 0.1, \sigma = 5</math></p>	Some predict huge success, others predict negative consequences.	Secret research <sup>1</sup>	Scrap idea

*Notes:* Rows represent different initial expectations about an innovation's effects,  $p(\pi)$ . Superscripts 1 and 2 identify innovation pathways. Pathway 1 treats row 1 as the baseline, and raises variance,  $\sigma$ . Pathway 2 considers shift from low to high  $k_D$  across columns.

The following expectation summarizes one pathway to research under the assumption that  $k_D$  is 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 whether the innovation will yield benefits or costs once deployed ( $e_0 \approx 0$ )—if she were confident that it would yield benefits  $e_0 > 0$ , she would pursue open research—and
- The improvement value ( $\theta$ ) is not too large, and there is little preexisting scientific research. Therefore, the researcher is not confident in her initial expectation ( $\sigma$  is high). If she were more confident that she understood the idea's effects ( $\sigma$  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 1 we showed that the manager pursues research only if her expected benefits for success are sufficiently high. Deep uncertainty means that an idea could generate

large positive, or negative, effects. When D weighs these different outcomes, his expectation of benefits is near zero. 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 (approximately zero).

In Proposition 1 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 result, when the costs and expected benefits are both low, the researcher is willing to pursue secret research *so long as she believes her research will convince the manager to approve her idea*—that is, when  $\sigma$  is high.

There are two reasons for this. First, when there is little preexisting research, the researcher's pilot research carries a larger weight in the manager's overall expectation of success. Second, when projects are likely to have either extreme positive or negative consequences, pilot testing indicates which direction the program will go. If 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 ( $k_D$ ) and the expected consequences of deploying a new innovation ( $e_0$ ). Substantively,  $k_D$  captures how sensitive the manager (and the institution at large) is to the expected moral and political costs associated with research when they authorize it.

**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 research costs is also high ( $k_D$  is high), but either the researcher's sensitivity is lower ( $k_R < k_D$ ) or cost sharing is moderately calibrated to support proposition 1.

If the manager's 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 facilitates 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 the unit share of research costs knowing that the manager will bear approval costs once research is complete.

### *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 from secret versus open institutions? We find that secrecy allows organizations to consider ideas that seem bizarre, morally

repugnant, or likely to fail when first conceived.<sup>60</sup> This leads to a straightforward expectation.

*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 willing to pursue these ideas only because the potential upside is high. The initial idea must have a large enough chance of making a positive impact for a researcher to pursue it. If research confirms that the idea is harmful, the institution scraps it early on. In the rare cases when research suggests that 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. Consider two comparable cases, a baseline case where R pursues open research because  $e_0$  and  $e_1$  are sufficiently large, and a counterfactual case where R pursues secret research because the counterfactual values  $e_{0\alpha}$  and  $e_{1\alpha}$  are smaller. Then so long as the true effect of innovation ( $\pi$ ) is large, increasing the true effect of innovation further increases the chance of innovation in the counterfactual case more than it does in the baseline case.*

There are two parts to this reasoning. First, in cases where managers approve open research, they are basically sold on the concept. Thus, even if the pilot shows only moderate success, they will approve innovation. By contrast, in cases where researchers opt for secrecy, the manager starts out skeptical. Thus, the result of the pilot tests must be very strong to convince the manager to approve innovation. Second, the pilot test is correlated with the true effect. Thus, increasing the true effect has a greater impact on whether the pilot's result will induce secret research. This has an interesting empirical analog. For every handful of bizarre and shameful failed projects, such as bionic cat robots or nuclear-induced tsunamis,<sup>61</sup> secret institutions provide a radical success—the reconnaissance satellite, for example. With foresight, these innovations all sounded risky. With hindsight, some are radical innovations that shaped the industrial and digital revolution and medical sciences.

### *Connecting the Mechanism to the Principal–Agent Literature*

The basic model identified how secret innovation allowed actors to distribute costs at different stages of the innovation process so they could pursue a wider range of novel ideas. However, the model did not fully explore the perverse incentives that arise

60. For a long list of failed innovations, see Houghton 2019.

61. Houghton 2019.

given uncertainty and principal–agent situations. We now introduce these 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.

### Monitoring

We assumed that if researchers exploit secrecy, the manager is forced to take on a cost  $k_D x$  when the program comes to light. In practice, managers can monitor subordinates' activities by asking for project details. Given that the researcher's actions can impose costs on the manager, why doesn't the manager monitor their activities?

This is at the heart of the principal–agent literature. Managers want to stop subordinates from taking actions they would not approve of. In this literature, there is agency loss because monitoring is difficult and expensive. However, if monitoring were costless and perfect, D would always monitor, and R would always behave.<sup>62</sup> This concern is relevant for our theory because R uses secrecy only because D will not approve.

We adjust the baseline model to capture monitoring as it is commonly studied in the principal–agent framework. First, we introduce uncertainty over research costs. We start with a simplifying assumption,  $k = k_R = k_D$ . We then add a step at the beginning of the game where Nature selects the cost associated with research,  $k \sim f()$ , where  $f()$  is supported on the nonnegative 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 that the researcher started a secret research program, he has two options: to allow it to continue or to shut it down. If he allows it to continue, the game reverts to open research (and associated payoffs). If he shuts it down, he avoids research costs entirely, the researcher incurs costs  $k_R$ , and the research has no effect (we do not realize  $\theta, m$ ).

We explicitly assume that D pays no cost to monitor, and if he does monitor, he perfectly observes R's behavior. Indeed, this is the exact condition the principal–agent literature suggests should drive complete monitoring. Define  $\bar{k} = \lambda[e_1 + \theta - c_{Rx}]$ , and  $\underline{k} = \lambda[e_1 + \theta - c_D]$ . Assume  $0 < \underline{k} < \bar{k}$ . Further define

$$\mathbb{E}[k | \text{sr, nor}] = \frac{\int_{\underline{k}}^{\bar{k}} kf(k)dk}{\int_{\underline{k}}^{\infty} f(k)dk}.$$

This represents the expected cost  $k$  that D will incur if he fails to monitor, at the moment he must decide whether to monitor or not, and given his expectation that secret research (sr) has happened, and he did not observe research (nor).

62. Eisenhardt 1989.

*Proposition 2: Don't-ask-don't-tell equilibrium. Suppose conditions 1, 2, and 3 can be satisfied for some  $k = k_R = k_D$ . Then in the model where D can perfectly monitor R, if*

$$\mathbb{E}[k|sr, nor] < \frac{\lambda[e_1 + \theta - c_D]}{x} \tag{4}$$

*then the following pure strategies are a perfect Bayesian equilibrium.*

- *D does not monitor if research is unobserved. D approves open innovation if  $k < \underline{k}$ . Regardless of how research occurs, D approves post-research innovation if  $\mathbb{E}[\pi m] > c_D - \theta$ . Off path, if D decides to monitor, D shuts down research with a cost profile  $k \geq \underline{k}$ , then does not approve innovation. Also off path, D rejects innovation absent research.*
- *R scraps the project if  $k > \bar{k}$ ; conducts open research if  $k < \underline{k}$ ; and conducts secret research otherwise.*

*Secrecy facilitates innovation if  $k \in [\underline{k}, \bar{k}]$ .*

See Appendix A.6. 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 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 that 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 zero. However, if he does not monitor, he can reduce his costs through plausible deniability.

In this equilibrium, there are research protocols that are so controversial the manager does worse by allowing research to continue even though he incurs only a share  $x$  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 has the following empirical implications.

*Expectation 3: Don't-ask-don't-tell. When managers are alerted that a researcher is secretly researching and does not want to share the details, they elect not to monitor because they suspect the program is controversial. Managers allow secret research to progress so they can retain plausible deniability.*

*Expectation 4: Telling implies shutdown. If managers observe controversial details of a research program that a researcher secretly pursued, they shut down the parts of the program they observe.*



### *Trust When the Researcher Can Fabricate Her Report*

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

We adjust the baseline model to understand whether 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. 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:  $m_R \rightarrow \mathbb{R}$ .<sup>63</sup> When research happens in secret, the manager observes only the report  $m_R$ . We say the research report is honest if  $m_R = m$  and dishonest otherwise. Third, we allow  $D$  to set the researcher's cost profile  $c_R$ ,  $k_R$ , which represents a manager's ability to assign projects to staff. In short, we want to know whether managers can find a researcher (1) who is willing to conduct secret research; (2) who is willing to write an honest report no matter the outcome of her pilot; and (3) who the manager will believe. Finally, we want to know whether the manager would like to employ a researcher who pursues secret research.

*Lemma 2: If conditions 1–3 and*

$$\lambda > \frac{k_D x}{e_1 + \theta - c_D} \quad (5)$$

*are satisfied, then the manager employs a researcher who is honest, trustworthy, and willing to conduct secret research.*

See Appendix A.7 for a technical statement of lemma 2 and proof. Lemma 2 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 works only if the institution employs unscrupulous patriots. The researcher who 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 ( $k_R \rightarrow 0$ ), but*
- *highly sensitive to the foreign policy costs associated with deploying a project ( $c_R = c_D/x$ ).*

63. Trivially, adding dishonesty costs makes honesty easier.

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 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's 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 R is less sensitive to the costs of deployment ( $c_R < c_D/x$ ) 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 actors 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 that R does not want projects approved 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.*

#### *External Ambiguity and Calibrating Cost Passing*

Because secrecy makes oversight hard, managers could sustain plausible deniability of the devilish details if the researcher briefed the manager informally. This would facilitate oversight, while offsetting the manager's expectation of incurring the increased costs from authorization should the controversial aspects ever be exposed. However, even if managers learn informally and approve passively, they are still more exposed to costs in expectation than if they learned nothing. For example, if a controversial experiment comes to light, an investigator may piece together the manager's knowledge from unusually long meetings with the research team, coded messages, or depositions of subordinates. Thus, at the time the manager is informally briefed, the decision to approve research must still factor in the cost of professional disgrace and criminal liability from involvement ( $k_D$ )—and the expectation of incurring these costs from the informal briefing (call it  $x + z < 1$ ). Here, the expectation is lower than if the manager had written a memo authorizing the experiments but higher than if they were truly ignorant ( $x$ ).

In Appendix A.8, we extend the model to account for these issues. We set up the model as a tough test for internal secrecy, because the researcher faces strong incentives to brief informally to pass on at least some costs, and we assert that if the

researcher does so it does not meet our definition of internal secrecy.<sup>64</sup> And yet, we still find that researchers will exploit internal secrecy (that is, not brief the manager at all) rather than provide informal briefings when the underlying cost parameters ( $k_i, c_i$ ) are high. What is more, we show that the option to brief informally raises the chance that research occurs beyond the baseline model. This illustrates how modeling other loose reporting requirements that internal secrecy facilitates expands the conditions under which innovation occurs.

### *Institutional Design*

In theory, even the president, the most senior member of the executive, answers to Congress (and the public). If abuse is possible, why does Congress tolerate the institutional arrangement described here? Why doesn't Congress design institutions that hold executives accountable even when they do not learn the details? It is hard to address this question empirically because internal secrecy is an enduring feature of national security institutions. In the US context, the National Security Act of 1947 handed the executive and national security agencies enormous power to sustain internal secrecy.<sup>65</sup> And this legislative framework survived reform debates that followed intelligence failures and executive abuse.

One possibility is that reform is hard and institutions are sticky. But in Appendix A.9, we adapt the monitoring model to provide a strategic explanation for Congressional inaction. We introduce a higher-order principal (Congress) who first sets  $x \in [0, 1]$  (the level of internal secrecy), and then the game unfolds following the monitoring model analyzed in Proposition 2, given that  $x$ . Our setup closely reflects two features of Congress's abilities and incentives expressed in historical debates over reform. First, Congress's main power to influence national security employees is by passing *ex ante*, rather than scandal-specific, laws about appropriate conduct for all future cases of secret research. This includes when managers are supposed to monitor their subordinates, when subordinates must report their activities, and so on. Then national security employees are confronted with specific scenarios (for example, the decision to pursue a particular idea) knowing the laws that govern their actions. Second, Congress is aware that internal secrecy is necessary to sustain external secrecy. Others have shown that greater oversight, or even greater sharing within the national security community, runs the risk that foreign agents will learn about sensitive operations.<sup>66</sup> Thus Congress knows that the higher it sets  $x$ , the more likely it is US rivals will discover secrets and exploit them.

64. There is still an indirect effect of internal secrecy, in that the manager's ability to offset costs comes from the fact that an unmodeled higher-order principal cannot observe informal manager–researcher interaction. We also discuss another model where informal briefings can sustain internal secrecy, which yields stronger results in favor of our theory.

65. Byrne 2014, xv.

66. Joseph, Poznansky, and Spaniel 2022.

We focus on conditions where, as shown in Proposition 2, if  $x$  is sufficiently low Congress induces the researcher and manager to engage in the behaviors described in the don't-ask-don't-tell equilibrium. But if Congress sets  $x$  higher, they induce the researcher to never engage in secret research, and we observe only uncontroversial research the manager directly approves.

We identify two strategic explanations for why Congress would tolerate secrecy and the possibility of abuse (set  $x$  low). The first closely reflects the don't-ask-don't-tell mechanism. Congress also desires welfare-enhancing innovations, and knows innovation is less likely if  $x$  is high. When the costs or risk of abuse are low relative to the foreign policy stakes, Congress prefers to tolerate a risk of abuse for the same reason the manager prefers not to monitor. Second, when the trade-off between internal and external secrecy is severe, Congress prefers to tolerate the risk of abuse to prevent foreign agents from discovering secrets. This second mechanism potentially explains the unique amount of internal secrecy in national security agencies. For example, there is little cost of leaking innovations in education policy, because they will not be exploited by rivals. Thus Congress has no incentive to write laws maximizing internal secrecy. Concerns over national security leaks can cause Congress to tolerate the risk of abuse from internal secrecy in national security institutions. We show that radical innovation is a convenient byproduct.

## Testing the Argument

We trace the logic of secret innovation in two cases: the search for mind control (MKULTRA) and the first reconnaissance satellite (CORONA). Table 2 summarizes the case parameters, which we substantiate in later sections. As a reminder, our theory identifies two pathways to secret research. MKULTRA fits our high-risk-high-reward pathway. Its moral repugnance generated enormous political costs during the research phase. But the promise of mind control was seen as very large. CORONA fits our lower-cost-but-high-variance pathway. The political costs of CORONA are smaller because they stemmed mainly from perceptions of wasteful spending. But so little was known about the atmosphere and satellite telemetry that researchers found it hard to predict its chance of success.

## Mind Control

In the late 1940s and early 1950s, US policymakers became convinced that the Soviet Union and the People's Republic of China had mastered mind control.<sup>67</sup> According to Richard Helms, a longtime CIA official who would go on to become director of the

67. Thomas 1989, 94; see also CIA 1956.

agency, “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.”<sup>68</sup>

**TABLE 2.** *Summary of case coding*

---

*MKULTRA*

---

Research team (R)	Sidney Gottlieb, Richard Lashbrook
Managers (D)	Allen Dulles, Richard Helms
Anticipated cost when research approved ( $k_i$ is high)	Research program involved experiments on US citizens without their knowledge or consent. If the project were exposed, those who knew these controversial aspects could face prosecution, professional disgrace, job loss, and so on. Whether exposed or not, actors could be psychologically sensitive to inflicting harm on innocent civilians.
Apportioning costs ( $x, 1$ )	Prosecution and professional disgrace were most likely for those who ordered controversial aspects. Psychological harm applies only to those who know.
Distribution of expected benefits ( $p(\pi)$ reflects Table 1, row 2)	Actors were cautiously optimistic that mind control could be achieved. They thought the possible national security benefits were substantial, and saw little chance an effective mind control policy would hold disutility.

---

*CORONA*

---

Research team (R)	USAF Major General Bernard Schriever, USAF Colonel Frederic Oder
Manager (D)	Secretary of the Air Force Donald Quarles
F Expected costs ( $k_i$ is low)	Executive may perceive those who authorize as undermining Eisenhower's space for peace policy, and may be perceived as wasteful stewards of public funds at a time of tight budgets.
Apportioning costs ( $x, 1$ )	Quarles knew he was closely monitored by Congress and would bear responsibility if he approved.
Distribution of expected benefits ( $p(\pi)$ reflects Table 1, row 3)	Many believed satellites were not viable. Some were concerned that if they worked they could cause escalation or crises. Others were hopeful.
Institutional quirk	While funding of CORONOA was internally open at USAF, the CIA offered an internally secretive funding vehicle. See text for details.

---

Policymakers were hopeful they could unlock the mysteries for themselves.<sup>69</sup> They believed mind control was “of the utmost importance ... [and] could mean the difference between the survival and extinction of the United States.”<sup>70</sup> A declassified memo from the early 1950s lists core aims: “A. Can accurate information be obtained from willing or unwilling individuals. B. Can Agency personnel ... 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 ... ?”<sup>71</sup>

68. Kinzer 2019, 54.

69. Select Committee on Intelligence 1977, 385.

70. Kinzer 2019, 49.

71. Redacted 1952, 1; see also CREST 2011, 2.

In 1950, the CIA conducted some ad hoc experiments code-named BLUEBIRD and ARTICHOKE.<sup>72</sup> Even these initial projects were handled outside normal oversight channels. A memo to the CIA director 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.”<sup>73</sup>

Within a few years, CIA director Allen Dulles decided to “intensify and systematize” the CIA’s efforts, and in April 1953 he authorized Sidney Gottlieb to establish MKULTRA.<sup>74</sup> Gottlieb was allowed to conduct experiments with virtually no oversight. Years of controversial experiments followed. Consistent with our assumptions, CIA managers granted this level of secrecy to researchers partly because of external threats. The Technical Services Division was awarded “exclusive control of the administration, records, and financial accountings of the program” owing to fear that “public disclosure of some aspects of MKULTRA activity could ... stimulate offensive and defensive action in this field on the part of foreign intelligence services.”<sup>75</sup>

While Dulles gave the research team broad authority to conduct experiments involving “chemical and biological materials capable of producing human behavioral and psychological changes,”<sup>76</sup> he and other managers<sup>77</sup> were not privy to the controversial details of how this research was performed.<sup>78</sup> Gottlieb secretly tested the effects of LSD on unwitting, nonvolunteer subjects. Under Operation Midnight Climax, sex workers lured unsuspecting American citizens to a safe house in San Francisco where CIA staff secretly dosed them with LSD and monitored them.<sup>79</sup> MKULTRA also involved experiments on prisoners overseas.<sup>80</sup> When the Church Committee reviewed MKULTRA years later, it was these research practices

72. McCroy 2006, 26–27; see also Streatfield 2007, 27.

73. CIA and Staff 1950, 1.

74. Kinzer 2019, 72–73.

75. Earman 1963, 2. They also worried that public disclosure “could induce serious adverse reaction in US public opinion” (2).

76. *Ibid.*, 1.

77. Richard Helms, who served as assistant deputy director for plans during critical periods, sat between Dulles and Gottlieb in the institutional hierarchy. We code him as a manager. The declassified record suggests he knew more than Dulles, but how much more is unclear. For example, in May 1953 Helms called LSD “dynamite” and said he “should be advised at all times when it was intended to use it.” But he appears not to have been aware of some of the most egregious experiments (Select Committee on Intelligence 1977, 395–96). Moreover, evidence of Helms advocating for unwitting testing is clearest in 1963—as the program was being shut down (394). As Kinzer 2019, 154 notes, “Only two officers—Gottlieb and Lashbrook—knew precisely what it was doing.” Ultimately, this is not especially consequential for our analysis. Our theory works in tiered institutions. The case would thus still fit if Helms was informed of some but not all of what Gottlieb did.

78. Those above Dulles knew even less. As the Church Committee noted, “There were no attempts to secure approval for the most controversial aspects of these programs from the executive branch or Congress.” Select Committee on Intelligence 1977, 394.

79. Kinzer 2019, 141–52; see also NBC 1977.

80. Kinzer 2019, 106.

that caused them to conclude 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.”<sup>81</sup>

As we will see, this firewall between managers and researchers meant that the latter, who oversaw the experiments, were at greatest risk for potential criminal prosecution and professional disgrace. CIA managers, who were ignorant of the most controversial aspects of MKULTRA, suffered fewer costs.

In summary, several features of this case fit our high-stakes pathway for secret innovation. It involves two primary actors: the MKULTRA research team (with Gottlieb at the center), and CIA management (the most senior for the majority of the time was Dulles). At the outset, Dulles knew that if MKULTRA succeeded, it would generate large benefits ( $e_0$  was high). However, he also knew the necessary research would be controversial ( $c_i$  was high).<sup>82</sup> Starting from this position, three facts about this case match the choices our model predicts. First, the CIA hand-selected Gottlieb to oversee MKULTRA. Second, Gottlieb judged that highly controversial human subjects research was necessary for MKULTRA. He could have discussed these research plans with managers but chose to keep these details secret. Third, Dulles had several opportunities to learn what Gottlieb was up to but never asked.

### *Why Was Gottlieb Chosen?*

Gottlieb was not an obvious pick to lead MKULTRA. Although he had experience in government laboratories as a chemist, he did not have an intelligence background. Why was an intelligence outsider selected to lead a high-stakes and intensely secret project? In the extension used to characterize Lemma 2, we argued that when researchers conduct scientific tests in secret, it is easy for them to give managers the mistaken impression that their novel idea is more effective than the research suggests. Anticipating this problem, the manager must carefully select an unscrupulous patriot: a researcher who is insensitive to whatever controversy it takes to complete a research program, but who shares the manager’s desire to field only projects that will advance national interests.

This is exactly how CIA managers saw Gottlieb and others on the Technical Services staff. The agency 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.”<sup>83</sup> The problem in Dulles’s view was that certain parts of the CIA “had shown no stomach for further work on

81. Select Committee on Intelligence 1977, 386.

82. To be clear, they did not know how controversial. The inspector general who audited MKULTRA similarly acknowledged these trade-offs. Streatfield 2007, 87.

83. Kinzer 2019, 47.

humans.” As Thomas notes, however, “the Agency’s Office of Technical Services Staff (TSS) had no such qualms ... They would have no reservations about testing ideas on unsuspecting subjects, especially in such a vitally important and urgent area as brainwashing.”<sup>84</sup>

Accounts of Gottlieb’s personality in particular are telling. According to Kinzer, “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.”<sup>85</sup> When he later testified before a Senate subcommittee about MKULTRA, Gottlieb used language we would expect from unscrupulous patriots: “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 ... There was a real possibility that potential enemies ... possessed capabilities in this field that we knew 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.”<sup>86</sup>

Of course, Gottlieb had incentives to cast himself as patriotic during an inquiry into his conduct. However, his behavior in the final years of MKULTRA also fit this personality profile. We show the patriotic researcher pursues her project only because she believes the science is viable. If she learns her research will fail to advance national security interests, she will quit, even if no one is stopping her. Consistent with this logic, one reason key parts of MKULTRA ended after nearly a decade of experimentation was Gottlieb’s realization 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.”<sup>87</sup> It would be odd for a researcher motivated by pride to publicly declare their work a failure.

### *Why Did Researchers Opt for Internal Secrecy?*

If our theory is correct, Gottlieb and his team exploited internal secrecy because they knew CIA managers would refuse to let them continue the most controversial experiments if they figured out what they were up to. Unfortunately, Gottlieb never explicitly articulated why he kept the most controversial details of experiments from his managers. But the context surrounding his actions is consistent with our logic in three ways.

First, the experiments he was engaged in, particularly the parts having to do with surreptitious testing of unwitting subjects, were extraordinarily controversial.

84. Thomas 1989, 98.

85. Kinzer 2019, 50.

86. *Ibid.*, 238.

87. *Ibid.*, 198.



According to the inspector general's report of 1963, "Research in the manipulation of human behavior is considered by many authorities in medicine and related fields to be professionally unethical, therefore the reputations of professional participants in the MKULTRA program are on occasion in jeopardy." It also states that "some MKULTRA activities raise questions of legality implicit in the original charter."<sup>88</sup> A memo from the late 1950s titled "Influencing Human Behavior" similarly notes that "some of the activities are considered to be professionally unethical and in some instances border on the illegal."<sup>89</sup> Because of this, "CIA officers felt it necessary to keep details of the project" extremely tightly guarded.<sup>90</sup>

Second, several CIA managers later stated they would have stopped MKULTRA if they had known its full extent. Dulles was reportedly interested in trying "everything the Communists could have done" but knew that "the risks for him and the Agency were enormous. If it ever became known that the United States government had funded what would be unprecedented clinical trials—ones beyond all ethical acceptability—it would most certainly lead to the sudden end of his remarkable and brilliant career."<sup>91</sup> This is likely why, as we will see in the next section, he was cut out of the loop of the precise details of MKULTRA. One senior CIA official who was "excluded from regular reviews of the project" was strongly opposed to MKULTRA—when he learned about it. According to one account, "it is possible that the project would have been terminated in 1957 if it had been called to his attention when he then served as Inspector General."<sup>92</sup>

Although less directly relevant given the timing, Stansfield Turner, who served as CIA director in the late 1970s, expressed similar reservations: "It is totally abhorrent to me to think of using a human being as a guinea pig ... 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 intelligence agencies should undertake."<sup>93</sup>

A final piece of evidence that internal secrecy facilitated Gottlieb's experiments is that once Congress got wind of MKULTRA and asked to review the program files, Gottlieb destroyed them on "the verbal order of then DCI Helms" rather than handing them over.<sup>94</sup> This impeded subsequent investigations into what had transpired.<sup>95</sup> Gottlieb and Helms purportedly felt that the experiments "might be 'midunderstood,'" leading them to order "that every scrap of paper relating to the brainwashing experiments be incinerated."<sup>96</sup>

88. Earman 1963, 1–2.

89. Quoted in Streatfield 2007, 86.

90. Select Committee on Intelligence 1977, 406.

91. Thomas 1989, 100.

92. Select Committee on Intelligence 1977, 409.

93. Kinzer 2019, 234.

94. Select Committee on Intelligence 1977, 389; see also CREST 2011.

95. Maret 2018, 29.

96. Streatfield 2007, 332.

*Managers Built the System So They Would Be in the Dark*

Our theory suggests that managers will embrace ignorance because they know that if they do not investigate they will incur a small cost as an ignorant bystander, but they may accrue a large gain from a successful innovation. On the other hand, if they investigate, they are faced with the choice of incurring a large cost or shutting down the program altogether. Three case features support this logic.

First, Dulles minimized his exposure to MKULTRA's details from the outset.<sup>97</sup> 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."<sup>98</sup> Dulles's 1953 memo states that "the nature of the research and the security considerations involved preclude handling the projects by means of the usual contractual arrangements."<sup>99</sup> According to one account, "Dulles ordered the Agency's book-keepers to pay the costs blindly on the signatures of Sid Gottlieb and Willis Gibbons, a former US Rubber executive who headed TSS."<sup>100</sup> Helms, who was one of the few senior officials to have reasonable insight into MKULTRA, "avoid[ed] oversight even by the CIA's director, because he 'felt it necessary to keep the details of the project restricted to an absolute minimum number of people'."<sup>101</sup> Richard Lashbrook, one of the senior scientists alongside Gottlieb, purportedly stated at one point that "what was actually signed-off on was not the same as the actual proposal, or actual detailed project."<sup>102</sup>

Second, CIA managers went to great lengths to avoid looking into MKULTRA. 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 afterward. The internal investigation that followed accused the TSS of "fail[ing] to observe normal and reasonable precautions." In response, 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'."<sup>103</sup> Ultimately, however, these were not formal reprimands, had no effect on advancement, and did not lead to a termination of the experiments.<sup>104</sup> Surprisingly, but consistent with our theory, even after investigators uncovered wrongdoing in the narrow experiments related to Olson, they did not expand their audit to MKULTRA broadly. One senior CIA official cautioned that a formal reprimand "would hinder 'the spirit of initiative and enthusiasm so necessary in our work'."<sup>105</sup>

97. Maret 2018, 47.

98. Kinzer 2019, 73.

99. Dulles 1953, 1.

100. Marks 1979, 57.

101. McCroy 2006, 28.

102. Quoted in Maret 2018.

103. Select Committee on Intelligence 1977, 398.

104. McCroy 2006, 30.

105. Marks 1979, 84.

Third, when MKULTRA 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 before Congress to testify. Years later, he was implicated in a variety of lawsuits by families of victims of MKULTRA. 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.”<sup>106</sup>

## Overhead Reconnaissance

Our second case examines the origins of the first US reconnaissance satellite, CORONA. We chose it for three reasons. First, it verifies that our argument extends beyond morally repugnant programs like MKULTRA to the costs and risks faced by many technical innovations. Second, reconnaissance satellites are a tough technological test of our theory because they are hard to keep secret, and because the research needs for cutting-edge experts across many scientific areas made openness attractive. Finally, there are historical quirks that provide a quasi-counterfactual test. CORONA occurred in a unique period in which the CIA was not widely known to be in the business of technical intelligence. Because of this, we know what *would have* happened if an open organization—that is, the Air Force, where it was originally pitched—was the only avenue for authorizing this bold innovation. There, it was rejected.

### *The Open Origins of CORONA*

Monitoring the Soviet Union was a pressing issue for policymakers in the early Cold War.<sup>107</sup> As the Soviets’ ability to thwart US reconnaissance tools advanced, concerns about the continued viability of the U-2 spy plane grew. US policymakers wanted a more reliable option.<sup>108</sup> Thus, some in the Air Force conceived of Weapons System 117L (the antecedent to CORONA).<sup>109</sup> Responsibility for it was placed in the Western Development Division, which was managing ballistic missile development. According to a declassified history, “WDD had been established with handpicked military personnel and with special reporting channels for expediting program decisions.”<sup>110</sup> They initially solicited design bids from cleared government contractors. Lockheed won a contract, but funding challenges loomed.<sup>111</sup>

106. Kinzer 2019, 256–57; see also Calabresi, Cabranes, and Heaney 1998.

107. May 1998, 21. US interest in military satellites emerged circa 1945. Wheelon 1998, 30.

108. Greer 1973, 3.

109. Brugioni 2010, 200.

110. Oder, Fitzpatrick, and Worthman 1988, 4.

111. Dienesch 2016, 129.

The institutional structure surrounding WS-117L was internally open. The secretary of the Air Force, Donald Quarles, “responded to news of the [Lockheed] contract award by ruling that neither mockups nor experimental vehicles should be built without his specific prior approval.”<sup>112</sup> In other words, the research team could not pursue pilot testing without alerting their manager. Moreover, although WS-117L was technically a classified project, presumably to keep information from the Soviet Union, “program details were reported to, and approved by, Congress.”<sup>113</sup>

From the perspective of Air Force managers, approving research on WS-117L presented low (but nonzero) political costs, but uncertain *expected* benefits. There was deep uncertainty about whether satellites were viable despite their enormous potential. According to a declassified history, “the technology to be embodied in the WS117L satellite was largely unproven; no satellite had even been orbited, and little was known of problems that might arise in a weightless, airless environment.” It also notes that Quarles “was not actively hostile to the satellite program as such, but had developed strong views about reliability and using low-risk technology.”<sup>114</sup> There was also concern about unanticipated escalation. On the costs side, Eisenhower was promoting the “space for peace” initiative, which “became a credo of US policy in 1955.”<sup>115</sup> Decision makers worried that if they authorized WS-117L they would be perceived as acting contrary to such commitments. Further, WS-117L was so novel that research into it could be perceived as wasteful. Quarles understood “the administration’s commitment to eliminate ‘noncritical’ defense expenditures.” Weighing these costs and benefits, and despite the desire of the WS-117L research team, he “found ample justification for his stubborn refusal to approve the start of a meaningful development program.”<sup>116</sup>

After it became clear that Air Force management would not adequately fund WS-117L, a plan was hatched to pursue it secretly. The project, conceived by Colonel Oder, was known as Second Story.<sup>117</sup> It had two prongs. First, it would be announced that WS-117L was being canceled and replaced with a scientific satellite overseen by the Air Force. This was a cover story. At the same time, the project would be covertly restarted and accelerated under the auspices of the CIA.<sup>118</sup> As noted, the CIA was just getting into the business of technical intelligence and thus was not an obvious choice to handle the project. This is likely why it did not originate there. Interestingly, however, a handful of the individuals involved with WS-117L were familiar with the Office of Science and Technology after working on the highly classified U-2 project.<sup>119</sup> Thus, the very fact that they proposed this option, which was outside

112. Oder, Fitzpatrick, and Worthman 1988, 5.

113. *Ibid.*, 14.

114. *Ibid.*, 6–7.

115. *Ibid.*, 5.

116. *Ibid.*, 6–7.

117. Dienesch 2016, 131–34.

118. Oder, Fitzpatrick, and Worthman 1988, 10.

119. Richelson 2002, 23.

the “‘normal’ development cycle,” is highly suggestive that internal secrecy was viewed, at least by the research team, as a way to advance a bold and risky innovation.<sup>120</sup>

Sputnik’s success in October 1957 took policymakers by surprise. While their earlier behavior was obviously not conditioned by an event that had not yet taken place, the Soviet Union’s success in space altered their thinking, including on the importance and feasibility of this technology.<sup>121</sup> As such, the post-Sputnik period is effectively a separate case and beyond our current scope. Moreover, policymakers’ emphasis on limiting many discussions to oral briefings “owing to the extreme sensitivity” of the project means that “there are few official records in the project files bearing dates between 5 December 1957 and 28 February 1958.”<sup>122</sup> Nevertheless, our theory points to several key elements of this period that are worth highlighting.

First, the strong desire for external secrecy—in this case, concealing CORONA from the Soviets—meant that the CIA’s ability to “maintain effective secrecy” was of paramount importance.<sup>123</sup> Second, efforts to preserve external secrecy resulted in deep internal secrecy, as evidenced by Eisenhower’s admonition that “only a handful of people should know anything at all about it.”<sup>124</sup> The fact that the CIA director was “the only US Government employee authorized to spend money without substantiating vouchers” is also notable in that it almost certainly helped prevent higher-order principals like Congress from interfering.<sup>125</sup> Eisenhower’s apparent decision to approve CORONA via “a handwritten note on the back of an envelope,” combined with the heavy emphasis on oral briefings, is also consistent with our mechanism focused on plausible deniability.<sup>126</sup>

## Conclusion

We have argued that secretive national security institutions are more innovative *because* they are secret. Secrecy is not equally valuable at every stage of innovation. Rather, it allows an enterprising researcher to pursue initial ideas that are so bizarre, morally controversial, or unlikely to work *ex ante* that their manager would refuse to fund the initial concept. 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 of the last century. The model explains that this theoretically drives different patterns of innovation in national security and other public-sector agencies.

120. Oder, Fitzpatrick, and Worthman 1988, 9. Initially, Second Story was “entirely concocted within Schriever’s own organization” (12).

121. Wheelon 1998, 32.

122. Oder, Fitzpatrick, and Worthman 1988, 15.

123. Brugioni 2010, 200–201. See also Wheelon 1998, 33. Consistent with our internal-secrecy logic, officials cite domestic concerns as one major justification for bringing testing to CIA.

124. Oder, Fitzpatrick, and Worthman 1988, 20.

125. *Ibid.*, 21.

126. *Ibid.*, 28. Some declassified CORONA documents still have redacted dollar amounts. CIA 1958, 5.

While we emphasized the welfare-enhancing effects of internal secrecy, our framework is general. Future researchers should explore the promise and peril of secrecy for innovation. They could consider how other institutional features could maximize innovation while reducing the risk of waste and abuse. They could also examine diversity in institutional design to harness the late-stage advantage of open organizations and the early-stage advantages of secrecy.

These insights also have significant policy implications, particularly with respect to the return of great power competition generally and competition between the US and China specifically. On the one hand, innovation is viewed as a key pillar of this dynamic. Mike Rogers and Glenn Nye, two former US representatives on opposite ends of the political spectrum, argued in an op-ed that “the race to take leadership in advanced technologies such as artificial intelligence, quantum computing, and 5G networks will determine the future balance of geopolitical power.”<sup>127</sup> On the other hand, officials have emphasized political and ideological factors as relevant to great power competition. The Biden administration’s National Security Strategy makes frequent mention of transparency and openness as being integral to competing with opaque, closed states like China and Russia.<sup>128</sup>

Our framework and findings suggest that there is a potential tension between these two impulses. In particular, internal secrecy—which sits uncomfortably alongside calls for greater openness domestically and internationally—has facilitated some of the most radical innovations of the last century. Ultimately, the best course of action may be to maintain a diversity of institutions.

## Supplementary Material

Supplementary material for this article is available at <<https://doi.org/10.1017/S0020818324000250>>.

## 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.
- 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.
- Biden, Joseph R. 2022. National Security Strategy. The White House.
- 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. *American Political Science Review* 110 (1):198–214.

127. Rogers and Nye 2019.

128. Biden 2022.

- Brugioni, Dino A. 2010. *Eyes in the Sky: Eisenhower, the CIA, and Cold War Aerial Espionage*. Naval Institute Press.
- Byrne, Malcolm. 2014. *Iran-Contra: Reagan's Scandal and the Unchecked Abuse of Presidential Power*. University Press of Kansas.
- Cain, Bruce E. 2015. *Democracy More or Less: America's Political Reform Quandary*. Cambridge University Press.
- Calabresi, Guido, Jose A. Cabranes, and Gerald W. Heaney. 1998. *Kronisch v. United States* (1998). United States Court of Appeals, Second Circuit. 150 F.3d 112 (2d Cir. 1998), Docket No. 97-6116.
- Carnegie, Allison. 2021. Secrecy in International Relations and Foreign Policy. *Annual Review of Political Science* 24:213–33.
- Carnegie, Allison, and Austin Carson. 2018. The Spotlight's Harsh Glare: Rethinking Publicity and International Order. *International Organization* 72 (3):627–57.
- Carson, Austin. 2018. *Secret Wars: Covert Conflict in International Politics*. Princeton University Press.
- CIA. 1956. *Brainwashing from a Psychological Viewpoint*. CIA CREST, CIA-RDP78-02646R000100100002-4.
- CIA. 1958. *Project Corona*. CIA CREST, CIA-RDP75B00514R000200050012-7.
- CIA. 1960. *Security Controls, 1953–1956*. CIA CREST, RDP72-00121A000100040001-9.
- CIA, Inspection, and Security Staff. 1950. *Project Bluebird*. CIA CREST, CIA-RDP83-01042R000800010003-1.
- Coe, Andrew, and Jane Vaynman. 2020. Why Arms Control Is So Rare. *American Political Science Review* 114 (2):342–55.
- Colaresi, Michael P. 2012. A Boom with Review: How Retrospective Oversight Increases the Foreign Policy Ability of Democracies. *American Journal of Political Science* 56 (3):671–89.
- Colaresi, Michael P. 2014. *Democracy Declassified: The Secrecy Dilemma in National Security*. Oxford University Press.
- CREST. 2011. *(U) Project MK-ULTRA*. CIA CREST, 06760269.
- Debs, Alexandre, and Nuno P. Monteiro. 2014. Known Unknowns: Power Shifts, Uncertainty, and War. *International Organization* 68 (1):1–31.
- 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):1247–65.
- Dienesch, Robert M. 2016. *Eyeing the Red Storm: Eisenhower and the First Attempt to Build a Spy Satellite*. University of Nebraska Press.
- 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):362–80.
- Drezner, Daniel W. 2019. Technological Change and International Relations. *International Relations* 33 (2):286–303.
- Dulles, Allen W. 1953. *Project MKULTRA: Extremely Sensitive Research and Development Program*. CIA CREST, C06767515.
- Early, Bryan R., and Erik Gartzke. 2021. Spying from Space: Reconnaissance Satellites and Interstate Disputes. *Journal of Conflict Resolution* 65 (9):1551–75.
- Earman, J. S. 1963. *Report of Inspection of MKULTRA*. CIA CREST, C06767515.
- Eisenhardt, Kathleen M. 1989. Agency Theory: An Assessment and Review. *Academy of Management Review* 14 (1):57–74.
- Farrell, Henry, and Abraham L. Newman. 2019. Weaponized Interdependence: How Global Economic Networks Shape State Coercion. *International Security* 44 (1):42–79.
- Goldfien, Michael, Michael Joseph, and Daniel Krcmaric. 2024. When Do Leader Backgrounds Matter? Evidence from the President's Daily Brief. *Conflict Management and Peace Science* 41 (4):414–37.
- Goldfien, Michael A., and Michael F. Joseph. 2023. Perceptions of Leadership Importance: Evidence from the CIA's President's Daily Brief. *Security Studies* 32 (2):205–238.
- Greer, Kenneth E. 1973. Corona. *Studies in Intelligence* 17:1–37.
- Grissom, Adam. 2006. The Future of Military Innovation Studies. *Journal of Strategic Studies* 29 (5): 905–934.



- Haines, Gerald K. 1999. Looking for a Rogue Elephant: The Pike Committee Investigations and the CIA. *Studies in Intelligence* 42 (5):81–92.
- Hawkins, Darren G., David A. Lake, Daniel L. Nielson, and Michael J. Tierney. 2006. *Delegation and Agency in International Organizations*. Cambridge University Press.
- Hinton, Henry L. 2001. *Observations on GAO Access to Information on CIA Programs and Activities*. Government Accountability Office.
- Horowitz, Michael C. 2020. Do Emerging Military Technologies Matter for International Politics? *Annual Review of Political Science* 23:385–400.
- Horowitz, Michael C., and Shira Pindyck. 2023. What Is a Military Innovation and Why It Matters. *Journal of Strategic Studies* 46 (1):85–114.
- Houghton, Vince. 2019. *Nuking the Moon: And Other Intelligence Schemes and Military Plots Left on the Drawing Board*. Penguin Books.
- Houston, David J. 2000. Public-Service Motivation: A Multivariate Test. *Journal of Public Administration Research and Theory* 10 (4):713–28.
- Ivan, Cristina, Irena Chiru, and Ruben Arcos. 2021. A Whole of Society Intelligence Approach: Critical Reassessment of the Tools and Means Used to Counter Information Warfare in the Digital Age. *Intelligence and National Security* 36 (4):495–511.
- Jacobsen, Anne M. 2015. *The Pentagon's Brain: An Uncensored History of DARPA, America's Top-Secret Military Research Agency*. Back Bay Books.
- Johnson, Loch K. 2022. *The Third Option: Covert Action and American Foreign Policy*. Oxford University Press.
- Jones, Derek C., Panu Kalmi, and Antti Kauhanen. 2006. Human Resource Management Policies and Productivity: New Evidence from an Econometric Case Study. *Oxford Review of Economic Policy* 22 (4):526–38.
- Joseph, Michael F. 2021. A Little Bit of Cheap Talk Is a Dangerous Thing: States Can Communicate Intentions Persuasively and Raise the Risk of War. *Journal of Politics* 83 (1):166–81.
- Joseph, Michael F. 2023. Do Different Coercive Strategies Help or Hurt Deterrence? *International Studies Quarterly* 67 (2).
- Joseph, Michael F., and Michael Poznansky. 2018. Media Technology, Covert Action, and the Politics of Exposure. *Journal of Peace Research* 55 (3):320–35.
- Joseph, Michael F., Michael Poznansky, and William Spaniel. 2022. Shooting the Messenger: The Challenge of National Security Whistleblowing. *Journal of Politics* 84 (2):846–60.
- 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):467–99.
- King, Nigel. 1990. Innovation at Work: The Research Literature. In *Innovation and Creativity at Work: Psychological and Organizational Strategies*, edited by M.A. West and J.L. Farr, 15–59. Wiley.
- Kinzer, Stephen. 2019. *Poisoner in Chief: Sidney Gottlieb and the CIA Search for Mind Control*. Henry Holt.
- Kollars, Nina. 2017. Genius and Mastery in Military Innovation. *Survival* 59 (2):125–38.
- Kopel, Michael, and Christian Riegler. 2008. Delegation in an R and D Game with Spillovers. *Contributions to Economic Analysis* 286:177–213.
- Kuo, Kendrick. 2021. Military Innovation and Technological Determinism: British and US Ways of Carrier Warfare, 1919–1945. *Journal of Global Security Studies* 6 (3):1–19.
- Kuo, Kendrick. 2022. Dangerous Changes: When Military Innovation Harms Combat Effectiveness. *International Security* 47 (2):48–87.
- Kurizaki, Shuhei. 2007. Efficient Secrecy: Public Versus Private Threats in Crisis Diplomacy. *American Political Science Review* 101 (3):543–58.
- Lai, Edwin L.-C., Raymond Riezman, and Ping Wang. 2009. Outsourcing of Innovation. *Economic Theory* 38 (3):485–515.
- Laird, Burgess. 2020. The Risks of Autonomous Weapons Systems for Crisis Stability and Conflict Escalation in Future U.S.-Russia Confrontations. RAND Blog.



- Laursen, Keld, and Nicolai J. Foss. 2003. New Human Resource Management Practices, Complementarities and the Impact on Innovation Performance. *Cambridge Journal of Economics* 27 (2):243–63.
- Lee, Caitlin. 2023. The Role of Culture in Military Innovation Studies: Lessons Learned from the US Air Force's Adoption of the Predator Drone, 1993–1997. *Journal of Strategic Studies* 46 (1):115–49.
- Macdonald, Julia M. 2015. Eisenhower's Scientists: Policy Entrepreneurs and the Test-Ban Debate 1954–1958. *Foreign Policy Analysis* 11 (1):1–21.
- Malis, Matthew. 2021. Conflict, Cooperation, and Delegated Diplomacy. *International Organization* 75 (4):1018–1057.
- Malis, Matthew. 2024. Foreign Policy Appointments. *International Organization*. <<https://doi.org/10.1017/S002081832400016X>>.
- Maret, Susan. 2018. Murky Projects and Uneven Information Policies: A Case Study of the Psychological Strategy Board. *Secrecy and Society* 1 (2):1–85.
- Marks, John. 1979. *The Search for the "Manchurian Candidate": The CIA and Mind Control*. Allen Lane.
- May, Ernest. 1998. Strategic Intelligence and US Security: The Contributions of CORONA. In *Eye in the Sky: The Story of the Corona Spy Satellites*, edited by Dwayne A. Day, John M. Logsdon, and Brian Latell, 21–28. Smithsonian Institution Press.
- McCroy, Alfred. 2006. *A Question of Torture: CIA Interrogation, from the Cold War to the War on Terror*. Henry Holt.
- Miller, Gary J. 2005. The Political Evolution of Principal-Agent Models. *Annual Review of Political Science* 8:203–225.
- Miller, Nicholas L. 2022. Learning to Predict Proliferation. *International Organization* 76 (2):487–507.
- NBC. 1977. An Interview with Admiral Turner. CIA CREST, CIA-RDP99-00498R000200150007-0.
- Neads, Alex, Theo Farrell, and David J. Galbreath. 2023. Evolving Towards Military Innovation: AI and the Australian Army. *Journal of Strategic Studies*. 1–30. <<https://doi.org/10.1080/01402390.2023.2200588>>.
- Oder, Frederic E. E., James C. Fitzpatrick, and Paul E. Worthman. 1988. *The Corona Story*. Center for the Study of National Reconnaissance.
- Posen, Barry. 1984. *The Sources of Military Doctrine: France, Britain, and Germany Between the World Wars*. Cornell University Press.
- Poznansky, Michael. 2020. *In the Shadow of International Law: Secrecy and Regime Change in the Postwar World*. Oxford University Press.
- Redacted. 1952. *Special Research, Bluebird*. CIA CREST, 0000140401.
- Richelson, Jeffrey T. 2002. *The Wizards of Langley: Inside the CIA's Directorate of Science and Technology*. Westview Press.
- Rogers, Mike, and Glenn Nye. 2019. Why America Must Boldly Win the Technological Race Against China. *The Hill*, 21 October.
- Rosen, Stephen Peter. 1988. New Ways of War: Understanding Military Innovation. *International Security* 13 (1):134–68.
- 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):727–35.
- Select Committee on Intelligence. 1977. *Project MK Ultra, the CIA's Program of Research in Behavioral Modification (Appendix A)*. US Government Printing Office. Available at <<https://www.intelligence.senate.gov/sites/default/files/hearings/95mkultra.pdf>>.
- Streatfield, Dominic. 2007. *Brainwash: The Secret History of Mind Control*. St. Martin's Press.
- Taylor, Mark Zachary. 2016. *The Politics of Innovation: Why Some Countries Are Better than Others at Science and Technology*. Oxford University Press.
- Thomas, Gordon. 1989. *Journey into Madness: The True Story of Secret CIA Mind Control and Medical Abuse*. Bantam Books.
- Vaynman, Jane. 2022. Better Monitoring and Better Spying: The Effects of Emerging Technology on Cooperation. *Texas National Security Review* 44 (4):34–56.
- West, Michael A., and Neil R. Anderson. 1996. Innovation in Top Management Teams. *Journal of Applied Psychology* 81 (6):680–93.

- Wheelon, Albert. 1998. CORONA: A Triumph of American Technology. In *Eye in the Sky: The Story of the Corona Spy Satellites*, edited by Dwayne A. Day, John M. Logsdon, and Brian Latell, 29–47. Smithsonian Institution Press.
- Wolford, Scott, Dan Reiter, and Clifford J. Carrubba. 2011. Information, Commitment, and War. *Journal of Conflict Resolution* 55 (4):556–79.
- 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.
- Zoghi, Cindy, Robert D. Mohr, and Peter B. Meyer. 2010. Workplace Organization and Innovation. *Canadian Journal of Economics* 43 (2):622–39.

## Authors

**Michael Joseph** is Assistant Professor in the Department of Political Science at the University of California, San Diego. He can be reached at [mjoseph@ucsd.edu](mailto:mjoseph@ucsd.edu).

**Michael Poznansky** is Associate Professor in the Strategic and Operational Research Department at the US Naval War College. He can be reached at [michael.poznansky@usnwc.edu](mailto:michael.poznansky@usnwc.edu).

## Acknowledgments

We thank Erik Gartzke, Ron Hassner, Federica Izzo, Kendrick Kuo, Matthew Malis, Rachel Metz, Michael Miller, Erik Sand, Jane Vaynman, the Strategic Multilayer Assessment Speaker Session, and the participants in the Center of Peace and Security Studies Workshop at UC San Diego for helpful feedback. Thanks to the men and women of the national security community for helpful conversations. We also thank Peter Rosendorff, Ashley Leeds, and the rest of the editorial team at IO as well as the substantive and technical reviewers for feedback that greatly improved the manuscript. The views expressed are the authors' own and do not represent those of the US Naval War College, the Department of the Navy, or any other organization of the US government.

## Key Words

Innovation; foreign policy; technology; principal-agent; secrecy; game theory; bureaucracy; CORONA; MKULTRA; intelligence

Date received: March 4, 2023; Date accepted: June 25, 2024