WORKING PAPER · NO. 2018-74

Rebel Capacity, Intelligence Gathering, and Combat Tactics

Konstantin Sonin and Austin L. Wright

SEPTEMBER 2022

Becker
Friedman
Institute

Political
Economics
Initiative

5757 S. University Ave.
Chicago, IL 60637
Main: 773.702.5599
bfi.uchicago.edu

Rebel Capacity, Intelligence Gathering, and Combat Tactics*

Konstantin Sonin and Austin L. Wright[†]

September 2, 2022

American Journal of Political Science

Running title: Rebel Capacity and Combat Tactics

Keywords: Civil war, insurgency, economic shocks

---

*For advice and comments, we thank Karen Alter, Christopher Blair, Luke Condra, Peter Deffebach, Quy-Toan Do, Kai Gehring, Garance Genicot, Michael Gilligan, Matthew Gudgeon, Amanda Kennard, Susan Hyde, Michael Kofoed, Daniel Krcmaric, Dorothy Kronick, Sarah Langlotz, Francois Libois, Nicola Limodio, Andrew Little, John Loeser, David Mansfield, Luis Martinez, Aila Matanock, Michaela Mattes, Marc Meredith, Gerard Padro í Miguel, Owen Ozier, Kyle Pizzey, Michael Povilus, Robert Powell, Martin Ravallion, Peter Rosendorff, Arturas Rozenas, Mehdi Shadmehr, Alexander Stephenson, David Van Dijcke, Oliver Vanden Eynde, Juan Vargas, Duncan Walker, Alex Weisiger, Liam Wren-Lewis, Andrew Zeitlin, and Yuri Zhukov and seminar participants at Columbia University, Georgetown University, New York University, Northwestern University, UC Berkeley, University of Pennsylvania, and The World Bank Research Group. Faculty at the United States Military Academy (West Point) provided exceptionally detailed feedback. This material is based in part upon work supported by the Pearson Institute for the Study and Resolution of Global Conflicts. We also thank Jarnickae Wilson, for his important contributions to early versions of this paper, and Andrew Shaver, with whom the military records were obtained. Eli Berman, Ethan Bueno de Mesquita, and Jacob Shapiro are owed a particular debt of gratitude for their support of this and related projects. María Ballesteros, Connie Guo, Lana Kugli, Michelle Liang, Mariya Milosh, Yonatan Litwin, Paul Soltys, and Julian Waugh provided excellent research assistance. All errors remain our own.

[†] Harris School of Public Policy, The University of Chicago, Chicago, IL 60637. Emails: `ksonin@uchicago.edu` and `austinlw@uchicago.edu`.

Abstract

Classic and modern theories of rebel warfare emphasize the role of resource endowments. We demonstrate that intelligence gathering, made possible by these endowments, plays a critical role in determining specifics of how rebels launch complex attacks against better-equipped government forces. We test implications of a theoretical model using highly detailed data about Afghan rebel attacks, insurgent-led spy networks, and counterinsurgent operations. Leveraging quasi-random variation in opium suitability, we find that improved rebel capacity is associated with (1) increased insurgent operations, (2) improved battlefield tactics through technological innovation, increased complexity, and attack clustering, and (3) increased effectiveness against security forces, especially harder targets. These results show that access to capital, coupled with intelligence gathering, meaningfully impacts how and where rebels fight.

Verification statement The data and materials required to verify the computational reproducibility of the results, procedures and analyses in this article are available on the American Journal of Political Science Dataverse within the Harvard Dataverse Network, at: https://doi.org/10.7910/DVN/4EZPKF.

Word count 9465

> Attack [the enemy] where he is unprepared, appear where you are not expected.
>
> ---
>
> Sun Tzu, “The Art of War”, 5th cent. BC

1 Introduction

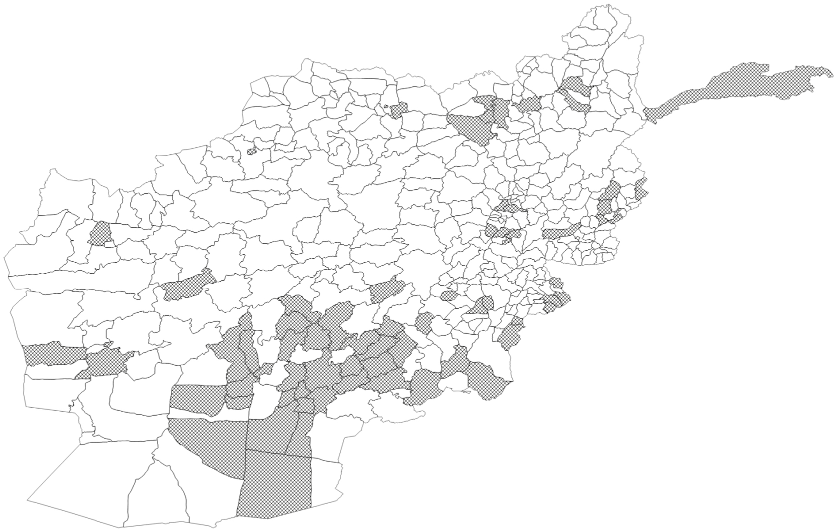
Intrastate conflicts have replaced interstate wars as the main source of human loss and population displacement. Generally, it is well-understood that resource endowments shape how rebels recruit, retain, and deploy their fighters (Weinstein, 2007). Fluctuations in rebel-held economic resources affect the scale of insurgent activity (Dube and Vargas, 2013), their control of strategic territory (Kalyvas, 2006), and how they treat civilians (Wood, 2014). These factors, in turn, impact whether civilians cooperate with rebels or collude with government forces (Condra and Shapiro, 2012*a*) and the ability of the government to engage in development and reconstruction (Sexton, 2016).

In this article, we unpack the impact of resource availability on combat tactics at a much more granular level. In our model of irregular warfare, rebels gather information about vulnerability of the targets and choose the pattern of attacks based on this information. Positive economic shocks enable rebels to acquire relatively high-quality intelligence and their attacks become more complex, involve more sophisticated weapons, and are clustered on a set of most vulnerable targets. Our model is a novel version of Colonel Blotto game, a standard general model of two-parties conflict (Blackett, 1958; Powell, 2007; Kovenock and Roberson, 2012). We add the possibility, for the attacking side, to gather additional information about targets' vulnerability: after the government has allocated its defense resources, each possible target is tested for vulnerability, and the rebels' choice of targets relies on results of these tests. In equilibrium, the optimal allocation of attacks across targets accounts for this additional information; this creates a link between the precision of information and allocation of attacks.

We test our model's implications using declassified military records provided by the U.S. government, which document hundreds of thousands of combat operations in Afghanistan during Operation Enduring Freedom. These data are unique in scope and scale; they encompass otherwise unobservable details about combat operations such as the location of insurgent surveillance operations, battlefield innovations by rebels, unit infiltration by insurgents, and use of deceptive weapon technologies. We combine these granular records with information on the location and intensity of microlevel opium production as well as satellite-derived measures of exogenous agronomic conditions that influence opium productivity. We leverage these high-resolution, high-frequency measures of agricultural inputs to construct a novel measure of exogenous opium suitability. We also gather a battery of additional information about agricultural price zones, infrastructure projects, irrigation technology, and use of coercive threats to manipulate local production to evaluate how shocks to rebel capacity impact rebel tactics.

We find consistent evidence that positive economic shocks to rebel organizations lead to an increase in violence and changes in *how* rebels produce violence. In particular, we find that increased rebel capacity is associated with more technological innovation by insurgents, additional attacks involving sophisticated infiltration of government forces, increased use of deceptive weapon technologies, and more complex, multi-target combat operations. We also find that rebels engage in more clustered attacks, both in time and space, as their access to capital grows. Importantly, since our combat records also include information about rebel-led surveillance, we can test the central conjecture of our theoretical model: access to more precise information about government vulnerabilities allows rebels to convert capital into more sophisticated attacks. Insurgents were able to conduct surveillance operations in 70 of Afghanistan's 398 districts at the start of our sample, which we visualize in Figure 1. We find broad evidence of this mechanism: violence and attack sophistication are particularly responsive to resource endowments in areas where insurgents conducted surveillance operations early in the campaign.

Figure 1: Rebel-Led Surveillance Operations Conducted in Afghanistan



Notes: Data on insurgent spy operations drawn from SIGACTS military records. Cross hatch pattern indicates insurgents conducted at least one detected surveillance operation during 2006, the first year of our sample. District boundaries are drawn from the ESOC Afghanistan map (398 districts).

A meaningful gap exists, as Berman and Matanock (2015) point out, between our understanding of *when* and *how* rebels engage in armed combat. Our paper helps to address this gap. Prior work provides compelling evidence that insurgents respond strategically to local economic shocks and aerial bombardment (Dube and Vargas, 2013; Berman et al., 2017; Vanden Eynde, 2018; Dell and Querubin, 2018), form alliances during war (Konig et al., 2017), and calibrate their use of violence against civilian populations (Condra and Shapiro, 2012*b*; Condra et al., 2018). Recent work also links exogenous economic shocks to terrorism financing and recruitment activity on the dark net (Limodio, 2019). We contribute to this literature by providing credible evidence that economic shocks to rebel organizations influence changes in combat tactics, which, in turn, has effect on battlefield efficiency.

More generally, our paper provides insights into the underlying mechanisms of insurgency. State capacity is central to economic theories of conflict (Besley and Persson, 2011; Powell, 2013). Yet the resources available to the state's competitors also influence when conflicts emerge, how internal wars are fought, and whether they end in withdrawal.

Our theoretical model and empirical tests also focus on a novel yet often overlooked dynamic of conflict: all sides collect information. Prior work on counterinsurgency has primarily studied how combat dynamics, including civilian harm, influence intelligence gathering by government forces (Condra and Shapiro, 2012*a*). This is, in part, due to the difficulty of observing how and when non-state actors engage in surveillance and manage the flow of information about combat activity and target vulnerabilities. Yet Kalyvas (2006) and others have noted, using various ethnographic, historical, and archival methods, the importance of information to all sides during conflict, especially asymmetric wars. Our theoretical model emphasizes this mechanism, suggesting that intelligence gathering shapes where, how, and to what effect violence is produced by rebels.

Our study is among the first to estimate the impact of resource endowments on battlefield effectiveness, notably attacks involving vehicle and weapons system damage as well as soldier casualties. We find that these attacks increase significantly with positive shocks to opium suitability, especially against hard targets. We also find that the impact of these shocks to rebel capacity is moderated by the intelligence gathering mechanism: increased combat effectiveness is sharpest in areas where insurgents have access to surveillance assets.

The rest of the paper is organized as follows. Section 2 introduces our theoretical model. Section 3 provides a brief overview of the institutional context. Section 4 details the empirical strategy. Section 5 presents the main results and robustness checks. The final section concludes.

> The spot where we intend to fight must not be made known; for then the enemy will have to prepare against a possible attack at several different points; and his forces being thus distributed in many directions, the numbers we shall have to face at any given point will be proportionately few.
>
> Sun Tzu, *ibid.*

2 Theory

In our model, the government chooses which potential targets to defend and the rebel group chooses the number of attacks using information it gathered about the targets' vulnerability. This is a Colonel Blotto-type model with asymmetric information that yields empirical implications we test using declassified military records from the recent Afghan conflict. We do not explicitly model the choice of weapons; it is relatively straightforward to demonstrate that an increase in endowments shifts the optimal choice towards the capital-intensive weapons, which is consistent with our empirical results. Instead, we focus on showing that improved intelligence results in more complex patterns of rebel attacks.

2.1 Setup

Consider a rebel group that attacks the government facilities using a certain technology (e.g., mortars). The group uses information of the quality $\theta \in \left[\frac{1}{2}, 1\right]$ to allocate the total of a attacks across different targets. The government optimally allocates resources to defeat attacks. This is visualized in Figure 2.

There are n potential targets for the government to protect. The government has resources to defend $r < n$ targets. Formally, the government strategy is a probability distribution $G(\cdot)$ over n-tuples $(g_1, ..., g_n)$ such that $\sum_{i \in n} g_i = r$ and $g_i \in \{0, 1\}$ for each $i \in n$. If an attack happens on the target that is defended, it does not succeed; if an attack is on an unprotected target, it succeeds with probability $p \in (0, 1)$, which parameterizes the quality of the attack technology. Since any deterministic choice of protection will result in rebels concentrating on unprotected targets, any reasonable placement of protection should be randomized.

After the government allocates protection, rebels gather intelligence about different targets' vulnerability. Specifically, rebels receive noisy signals $(s_i)_{i \in n} \in \{0, 1\}^n$ that are determined according to

$$P(s_i = 0 | g_i = 0) = P(s_i = 1 | g_i = 1) = \theta.$$

We assume that the signals are informative: $\theta > \frac{1}{2}$, i.e., a target that is unprotected is more likely, than not, to be marked "vulnerable" ($s_i = 0$). A higher θ results in more informative signals.

The rebels' strategy is a mapping $F(a_1, ..., a_n; \cdot)$ from the n-tuples of signals about periods' vulnerability $(s_1, ..., s_n)$ into a probability distribution on n-tuples $(a_1, ..., a_n)$ such that $\sum_{i \in n} a_i = a$ and a_i is a non-negative integer for each $i \in n$. The possibility to base the attacking strategy on additional information about the vulnerability of targets is novel in the Colonel Blotto games and, more generally, formal conflict literature. If $\theta = \frac{1}{2}$, then the signals are totally uninformative and the game is a standard Colonel Blotto game with n targets (Kovenock and Roberson, 2012). If $\theta > \frac{1}{2}$, then for each realization of signals $(s_1, ..., s_n)$, the attacker's optimization problem is similar to that in Powell (2007), in which targets have heterogeneous values rather than heterogeneous probabilities of being vulnerable.

The rebel group maximizes the probability of at least one successful attack. The government is interested in minimizing this probability.

The game starts with the government allocating its defensive resources across n targets. Then rebels receive noisy signal $(s_i)_{i \in n} \in \{1, 0\}^n$ about the government's defense and choose the distribution of their attacks across the targets.

Definition 1 *Given the resources available to rebels, an equilibrium is rebels' choice of a c.d.f.* $F^*(s_1, ..., s_n; \cdot)$, *which is a function of signals about each target, into a probability distribution over* a *attacks on each of the* n *targets, and the government's choice of a c.d.f.* G^* *over* $r-$*combinations of* n *targets. Given* G^*, F^* *maximizes the probability of at least one successful attack; given* F^*, G^* *minimizes it.*

2.2 The Attacking Game

We start backwards. It is straightforward to establish that the government allocates resources into r targets chosen randomly and uniformly across all possible combinations. The rebels' optimal strategy depends on the signals that they observe. Information gathering results in x "vulnerable" ($s_i = 0$) and $n - x$ "defended" ($s_i = 1$) targets, where x is a random outcome. The rebel's assessment that a particular target i is vulnerable is based on two pieces of information: first, the signal about its vulnerability, s_i, and the total number of "vulnerable" signals, $x = \#\{j|s_j = 0\}$, which is a random variable, the sum of two random variables with binomial distributions with different probabilities of success: $n - r$ vulnerable targets produce signal 0 with probability θ, while r defended targets produce signal 0 with probability $1 - \theta$.

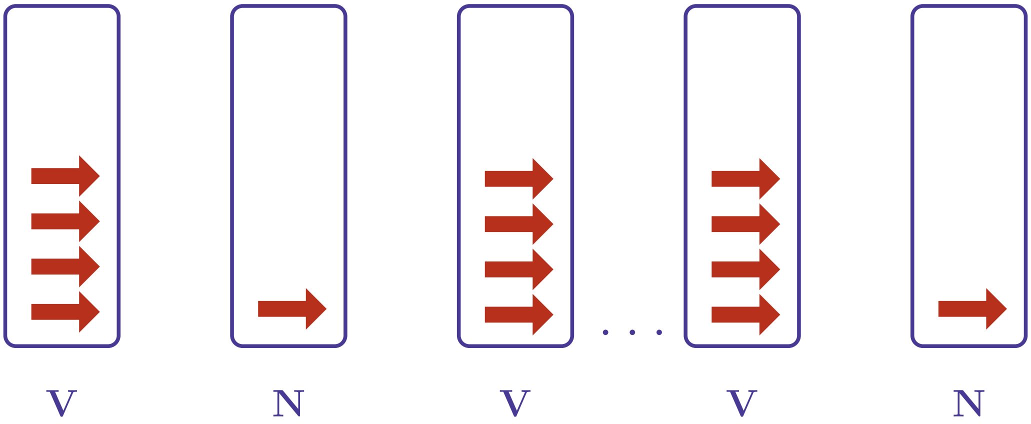
Intuitively, when resources are scarce, a signal about vulnerability of one target is informative about the vulnerability of other targets, even though signals are conditionally independent. The number of "vulnerable" signals x can be any integer between 0 and n with a non-zero probability. In the two extreme cases, $x = 0$ (intelligence signals that no target is vulnerable) and $x = n$ (all are "vulnerable"), there is no information to update upon. In all other cases, $1 \leq x \leq n - 1$, signals are informative:

$$P(g_i = 0|s_i = 0) > P(g_i = 0|s_i = 1).$$

Consider the rebels' choice of one attack across two targets with conditional (on x) probabilities of being vulnerable q_1 and q_2, respectively, with $q_1 > q_2$. If there are no attacks already planned on these targets, then an attack timed on the first one provides a higher marginal probability of success. Yet the fact that the vulnerable target has a higher probability of success does not mean that all attacks should be concentrated on it. Indeed, suppose that there are already m attacks launched on the first target, $m \geq 1$, and no attacks launched on the second. One more attack on the first target results in $q_1 (1 - p)^m$ of marginal probability of success. An attack on the second one contributes $q_2 p$. Thus, given sufficient capacity, it is optimal to launch some attacks on the second, less likely to be vulnerable, target.

The rebels' optimal strategy is determined by $\rho_{n,r}(x|\theta)$, the ratio of the probability that the target is vulnerable to the probability that the target is defended, both probabilities

Figure 2: Optimal Rebels' Strategy in the Attacking Game



Notes: After rebels recieve information about the targets' vulnerability, the optimal allocation of attacks across targets requires allocating more attacks on targets that are labeled "vulnerable" (V), and fewer attacks on other targets (N).

conditional on the total number of "vulnerable" signals x. As demonstrated in the appendix (see page A-6 of the online appendix), this ratio can be derived using the following recursive formula:

$$\rho_{n,r}(x|\theta) = \frac{n - x + 1 + (x-1)\rho_{n-1,r}(x-1|\theta)}{\frac{n-x}{\rho_{n-1,r}(x|\theta)} + x} \text{ for } 1 \leq r,\ x \leq n-1,$$

with the recursion on both total number of targets, n, and the total number of "vulnerable" signals received, x. Using the vector of critical ratios $\rho_{n,r}(x|\theta)$, $1 \leq x \leq n$ one can demonstrate that the optimal strategy requires, for each x, allocating the first attacks against the targets with "vulnerable" markers until the threshold $d_{n,r}(x|\theta) = \min_{i\geq 1}\left\{i|\rho_{n,r}(x|\theta)\,(1-p)^i < 1\right\}$ for each "vulnerable" target is reached, the next n attacks against targets with "defended" signals, then against the "vulnerable" targets again, etc. Proposition 1 states the result formally.

Proposition 1 *There exists a unique equilibrium in the attacking game. The government protects r targets chosen randomly and uniformly across all possible combinations and rebels follow the signals that they receive. For any number x of targets that are "vulnerable" (have $s_i = 0$), there is an optimal number of attacks $\overline{a}(x)$ such that $\min\{a, x\overline{a}(x)\}$ attacks are* *distributed uniformly over* x *"vulnerable" targets. The remaining* $a - \min\{a, x\overline{a}(x)\}$ *attacks are distributed uniformly across* $n - x$ *"defended"targets.*

2.3 The Role of Intelligence

The rebels' equilibrium strategy described in Proposition 1 depends on the quality of information parameterized by θ. Specifically, higher precision of information leads to a higher clustering of attacks: more attacks are launched on targets that intelligence gathering identified as vulnerable. In other words, the rebels' attacking strategy becomes more complex with more intelligence gathering. Mathematically, the information entropy, $\sum_{i=1}^{n} \frac{a_i}{a} \ln \frac{a_i}{a}$, a standard measure of complexity, is at its maximum when attacks are purely random, i.e., distributed uniformly across targets, and goes down, when attacks become more sophisticated. In our case, an increase in complexity follows an improvement in intelligence gathering, which is consistent with the empirical results of Section 5. Proposition 2 provides a formal result.

Proposition 2 *For any amount of rebels' resources, the higher is the precision of information that rebels receive,* θ*, the higher is the clustering (concentration) of attacks, i.e., the lower is the expected number of unique targets attacked and the larger is the expected number of attacks, both total and successful, per target attacked.*

The critical element of Proposition 2 is that for any number x of "vulnerable" targets, the probability $q(x)$ that a target marked "vulnerable" is indeed vulnerable is (weakly) increasing in the precision of information θ, and thus the critical threshold $\overline{a}(x)$ is (weakly) increasing in θ for any x. As a consequence, more precise information leads to a higher clustering of attacks: more attacks are launched on a smaller number of targets.

In equilibrium, the optimal choice of rebels depends, in addition to precision of information θ, on the number of potential targets for attacks n, the resources in the disposal of the government r, and the efficiency of weapons p. Proposition 2 establishes that a higher precision of information, e.g., as a result of an increase in revenues, leads to attacks becoming more sophisticated and, naturally, more effective. In line with the theoretical model, in

Figure A-1 and Table A-2 (see pages A-15 and A-16 of the online appendix, respectively), we present descriptive evidence indicating a robust association between exogenous variation in opium suitability and our benchmark measure of surveillance activity. More generally, the increase in complexity (mathematically, a decrease of entropy) as a result of an increase in resources is one of the central empirical implications tested in Sections 4-5.

The comparative statics results with respect to the government's resources and rebels' weapons efficiency are intuitive. An increase in government resources, a higher r, decreases the probability of rebels' success as the *ex ante* probability that each target is protected increases. With a lower probability that each target is vulnerable, the marginal return to an increase in complexity (or to a move to a more sophisticated weapon) is lower. This has the same effect as the fall in rebels' revenues, which results in less information gathering and, therefore, less sophisticated attacks.

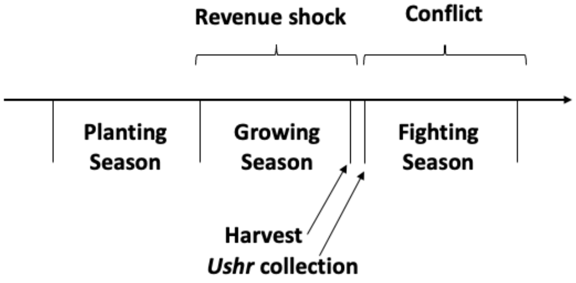
The theoretical model can be extended in a number of ways. At the expense of much cumbersome algebra, it is possible to incorporate the trade-off between an increase in the number of attacks and an increase in the precision of intelligence gathering as a result of an increase in rebels' resources. For our purposes, it is sufficient that an increase in rebels' resources cannot lead to a decrease in the quality of information. It is also possible, perhaps more realistically, to model a dynamic, rather than a one-shot interaction between the government and rebels. Given that our own estimates demonstrate the very limited ability of rebels to smooth the availability of their resources over fighting seasons (see Table A-3 on page A-17 of the online appendix), we leave the dynamic extension for future papers.

3 Institutional Context: Afghanistan

The literature on Afghanistan conflict is vast. Here, we briefly highlight several important dimensions: timeline; organization; opium cultivation and trafficking; and, conflict dynamics. Prior to the US-led invasion in 2001, the Taliban had held partial control of government operations since September 1996, when they seized Kabul from President Burhanuddin Rabbani. Following the Taliban's removal from power, Hamid Karzai was appointed as president and re-elected in elections in 2004 and 2009. Forces representing the North Atlantic Treaty Organization (NATO) and coalition partners, including large detachments from the United States, supported Afghan reconstruction and provided security support to the Afghan army and, eventually, national and local police. While these forces gradually expanded their presence across the country, the Taliban regrouped in Pakistan. In 2006, the Taliban's insurgency began in force, engaging in attacks across the country. With violence on the rise and Afghanistan's emerging democratic institutions at risk, a surge of coalition forces was authorized in 2009, with the international troop presence reaching a peak in 2011. In March 2011, Karzai announced the first of five tranches of the security transfer from international to local forces. By the end of 2014, Operation Enduring Freedom (OEF) had concluded with the final International Security Assistance Force (ISAF) transition ceremony and shift to Operation Resolution Support. In this article, we study the period from 2006, when the Taliban's insurgency emerged, to 2014, with the end of OEF and the ISAF mission.

The Taliban also encouraged the production of opium during the study period, in a sharp contrast with the last years in power before the 2001 US-led invasion. By 1999, after three years in power, the Taliban were largely isolated from the international community (Farrell and Thorne, 2005). Islamic fundamentalism, including concerns about the treatment of women and ethnic minorities, coupled with concerns about the production and export of nearly 70% of the global supply of illegal opium were key to this political isolation. In exchange for extensive diplomatic engagement with the UNODC and promises of developmental aid from the United Nations more broadly, the Taliban announced a large-scale counternarcotics and eradication problem. Mullah Omar, leader of the Taliban at the time, issued a $fatwa$ (religious decree) banning the production of opium. Continued cultivation risked public humiliation or execution. Production declined rapidly and farmgate prices increased significantly. This large-scale economic shock may have weakened Taliban control in opium producing areas and beyond.

Figure 3: Seasonality of Revenue Extraction and Conflict in Afghanistan



Notes: various seasons visualized drawing on UNODC descriptions of production cycles and seasonality of conflict present in combat records. Planting occurs in late fall and early winter; growing season primarily ranges from February to April; harvest and tax collection is conducted in April, May, and June; fighting season runs until late September.

Once the Taliban were removed from power in 2001, they recognized that taxes collected from opium farmers and protection payments from traffickers could support their war efforts and quasi-state public goods provision (Felbab-Brown, 2006). By 2006, the beginning of our study, the UNODC estimates that more than half of the Afghan GDP was tied to the drug trade. Afghan opium production reached a record high in 2014, the final year of our study, with estimated production exceeding 210,000 hectares for the first time.

We visualize the opium planting and growing and fighting seasons in Figure 3. In the primary opium producing regions, seeds are planted in late fall and early winter. The growing season typically ranges from February to April, with most opium latex harvested and packaged in April, May, and June. Peters (2009) provides a thorough review of the industrial organization of the Taliban. Taliban commanders and veteran fighters return from Pakistan in June to collect taxes from opium farmers (*ushr*, typically a flat 10% fee mandated by the Quran). Taxes can be paid in currency, opium blocks, or other goods, such as motorcycles, offroad vehicles, and weaponry. The Taliban also benefits from protection fees levied on opium traffickers as they pass through rebel-held territories. Opium farmers, refiners, and traffickers rely on the Taliban to sustain a political and security environment and ensure that cultivation, refinement, and exports continue with minimal government interference (Na, 2018; Giustozzi, 2019).[1]

Taxes are collected by fighters and receipts are distributed to farmers to prevent double taxation. Fighters pass their collections to district-level commanders (equivalent in scale to US counties). Taxes are subsequently passed upward to provincial and regional commanders, who keep ledgers of their annual revenue and are subject to audit by the Taliban's Central Finance Committee, based out of southwestern Pakistan. Most proceeds remain with the district commander, for conducting operations in the subsequent fighting season which typically lasts until September. These funds can be used to purchase weapons and ammunition, as well as covering the salaries of fighters and rebel informants. The Central Finance Committee (CFC) retains the authority to demote or relocate field commanders to less desirable fronts if audit irregularities are found. The remaining revenue is split between supporting operations conducted in resource-poor districts where local taxes alone are insufficient for supporting rebel attacks and developing Taliban infrastructure in Pakistan (including small-scale hospitals for wounded fighters).

We focus primarily on the period from 2006 to 2014. The industrial organization of the insurgency, most notably the taxation and command structures oriented around administrative districts, which are central to our argument that revenue influences combat tactics, emerged in 2006. Our military records track insurgent operations until the end of 2014, when the NATO Operation Enduring Freedom was transitioned to Mission Resolute Support.

4 Empirical Design

In this section, we review our microdata, and introduce our identification strategy.

4.1 Conflict Microdata

We exploit declassified records of the U.S. Department of Defense in 2006 to 2014, which catalogue combat engagements and counterinsurgent operations during Operation Enduring

---

[1] We discuss additional Taliban income sources in the Online Appendix (see page A-10 of the online appendix).

Freedom in Afghanistan (Shaver and Wright, 2016). The data platform was populated using highly detailed combat reports logged by NATO-affiliated troops as well as host nation forces (Afghan military and police forces). Data of this type differ substantially in coverage and precision from media-based collection efforts. For example, the combat records we study include information about the timing of any given attack, usually accurate to within minutes. In addition, our records include georeferences that are derived from satellite-linked devices that were deployed in the field rather than georeferencing of landmarks mentioned in journalistic coverage. As Weidmann (2016) notes, this compilation of conflict events is the most complete catalogue of combat engagements during the war in Afghanistan.

These data include information about a number of types of violence, including direct fire engagements, indirect fire events, and IED explosions. Direct fire attacks are primarily line-of-sight, close combat events. Indirect fire consists of mortars and other weapons that can be deployed without close contact with military forces. IEDs consist of explosives that have been emplaced and are detonated through a variety of trigger mechanisms (pressure plate, cable-to-battery, radio signal, laser beam, etc.). For each event, we can track the target or targets involved as well as the outcome of the event (whether an attack caused damage or casualties). We also observe information about when coalition forces engaged in search and seize operations, gathering potentially actionable information about insurgent operations, as well as insurgent detentions. The records include additional information about insurgent activity that could influence civilian involvement in opium production. In particular, insurgents may use violent and non-violent tactics to intimidate civilians, such as killings of government collaborators and the posting of 'night letters' and other non-lethal shows of force. Our military records include information about these tactics as well, enabling us to address potential concerns about rebel involvement in local opium production.

In addition, we have information on several novel, previously unreleased dimensions of combat. First, soldiers on deployment were told to document instances of tactical innovation by insurgents. These tactics and procedures reports identify instances of rebels engaging in new attack formations, focusing on new targets, developing and/or deploying new weapons technology, or otherwise adopting an unexpected improvement in their combat operations. Although the original text files attached to each event would have provided extensive details about what specific innovation was observed for each report, this information was removed from authors' access during declassification. On its own, however, this measure serves as an indicator of technological innovation in the battlefield. Second, our combat records include information about false or hoax explosive devices deployed by insurgents. These events indicate where and when rebels are engaging in active deception as a tactic. There are several purposes for this type of deception. One is to learn about coalition movement, counter-IED technologies, or counterinsurgent activities. Another potential purpose is to pull government soldiers and assets towards one area while insurgents engage in a counter-maneuver, taking advantage of depleted forces in another area. Although we cannot deduce the motive for use of these deceptive tactics, we infer that the use of fake devices reveals a degree of technological sophistication. Third, our data contain a detailed record of insurgent-led insider attacks. These events occur when insurgents have infiltrated one or more Afghan security units, 'turning' members of the security forces against fellow members of the Afghan army and, in other cases, coalition forces. These attacks require cultivation of an insider or the recruitment of individuals to serve as double-agents, joining the armed forces with the intent of harming other members of the security force. These events are also uniquely disruptive to collaborative patrol operations as insider attacks typically led to segregation of forces and diminished joint operations. Fourth, the depth of our records enables us to identify complex events involving multiple targets. These events can involve, for example, an attack on infrastructure (e.g., a base), a force in transit to the infrastructure (e.g., a patrol returning to base), and government countermeasures (e.g., surface-to-air engagement). These complex events potentially suggest a high degree of coordination and fighting capacity. Finally, our records include information about rebel-led surveillance operations detected by counterinsurgents. This measure captures instances of suspected insurgents monitoring troop movements in and out of base locations, changes of guard or other base-specific activities, and patrol movement. These efforts to monitor troop and base activity are likely attempts to identify target vulnerabilities for later exploitation in combat.

4.2 Opium Cultivation, Suitability, and Prices

Opium production estimates are derived from ground-validated remote sensing techniques, which use high resolution satellite imagery to track changes in vegetation during the spring harvest. UNODC-Afghanistan randomly spatially samples potential agricultural zones within provinces and acquires pre-harvest and post-harvest imagery (see Figure A-2 on page A-15 of the online appendix, panel (a)). These images are then examined for changes in vegetative signatures consistent with the volatile wetness of opium plants after lancing. From this sampling technique, officers estimate the spatial risk and calculate granular estimates of opium production (see Figure A-2 on page A-15 of the online appendix, panel (b)). These gridded estimates are then compiled as the annual amount of opium production (in hectares) for each district. We correct for changes in the administrative boundaries of districts over time using the Empirical Studies of Conflict (ESOC) administrative shapefile. To translate production into yields, we compile additional details about annual yield (kilograms per hectare) from UNODC-Afghanistan annual reports.

To measure opium suitability, we gather climatic data, daily, district-level temperature (Kelvin) and precipitation (mm) measures from the National Centers for Environmental Prediction (NCEP) and the Department of Energy, which prepared the baseline climate reanalysis by using state-of-the-art assimilation techniques (see Saha et al., 2010, for full details). We construct parameters capturing the number of days within each growing season these data fall within a particular set of binned ranges, which enables us to account for non-linear relationships between weather conditions and agricultural productivity (Dell, Jones and Olken, 2014). We supplement these data with information from Food and Agriculture Organization's Harmonized World Soil Database, extracted using the district-level cross section. We include nutrient availability, nutrient retention, rooting conditions, oxygen availability, excess soil salts, toxicity, and packedness and workability (which impacts the ability to manage fields). For each district, we calculate the percentage of land mass where these soil features present no or slight limitations to productivity (Class 1 under the FAO guidelines). Because various combinations of weather and soil conditions may produce high and low productivity zones in a complex system, we interact these measures with our degree-day and precipitation-day measures. We merge these data with our panel data on opium production and produce a standardized fitted value of opium productivity given these exogenous parameters. We use the least squares estimation equation below.

$$
\begin{split}
ln(production_{d,t}+1) = \alpha + \sum_{i=1}^{7}(\vartheta_i Precip-Day_{d,t}) + \sum_{i=1}^{7}(\zeta_i Precip-Day_{d,t}^2) \\
+ \sum_{i=1}^{10}(\eta_i Temp-Day_{d,t}) + \sum_{i=1}^{10}(\rho_i Temp-Day_{d,t}^2) + \sum_{i=1}^{7}(\mu_i SoilQual_d) \\
+ \tau_{ij}\sum_{i=1}^{7}(Precip_{d,t}) \times \sum_{j=1}^{7}(SoilQual_d) + \phi_{ij}\sum_{i=1}^{10}(Temp-Day_{d,t}) \times \sum_{j=1}^{7}(SoilQual_d) \\
+ \gamma X_y + \varepsilon_d
\end{split}
\tag{1}
$$

where $log(production_{d,t}+1)$ is the production (log) for a given district, d, and growing season, y (year). X_y captures growing season fixed effects. $Precip-Day_{d,t}$ and $Temp-Day_{d,t}$ capture the effect of our precipitation-day and degree-day (temperature-day) parameters. We also include the square of these counts. $SoilQual_d$ captures the soil quality features noted above. We then fully interact these base terms. From this regression, we produce $ln(\widehat{production}_{d,t}+1)$, which is our unstandardized fitted value. Denote this value as $\Lambda_{d,t}$. We standardize this value using the following expression:

$$Suitability_{d,t} = \frac{\Lambda_{d,t} - \bar{\Lambda}_{d,t}}{var(\Lambda_t)^{-1}} \tag{2}$$

$Suitability_{d,t}$ is demeaned and standardized with respect to the standard deviation of the fitted values. This approach is most similar to Mejía and Restrepo (2014), who use land features and soil characteristics to predict coca production in Colombia. The primary difference between our two methods is the use of high frequency climatic inputs as well as the use of interactions to capture heterogeneous climatic effects via soil quality conditions.

Opium price data are compiled at national and regional levels. We rely on UNODC-Afghanistan documentation to assign districts to price zones. Although Afghanistan's aggregate opium exports represent more than 75% of global exports, only a small subset of district-years (.1%) reaches the price-maker threshold set in Bazzi and Blattman (2014) (10% of global exports). In addition, no district is a potential price-maker in our sample for more than half of the sample years. This suggests that nearly all districts are price-takers for nearly (or) all of the study period. Because of this, use of the aggregate price to calculate revenue is unlikely to be substantively biased. However, given the data available, we can implement an alternative supply-driven approach to price variation. Following UNODC reports, we find that aggregate, country-level production in the prior year is a primary driver of year-over-year variation in prices. We denote this quantity as $AggProd_{t-1}$. Naturally, increased aggregate production from the prior year drives down national prices in the subsequent year. Leveraging aggregate production yields plausibly exogenous variation in the price component of revenue (once we invert the value).

4.3 Empirical Strategy

We study the relationship between rebel capacity and violence, leveraging plausibly exogenous variation in opium suitability. Our baseline sample is a balanced panel of district-years from 2006 to 2014. We estimate the following OLS regression:

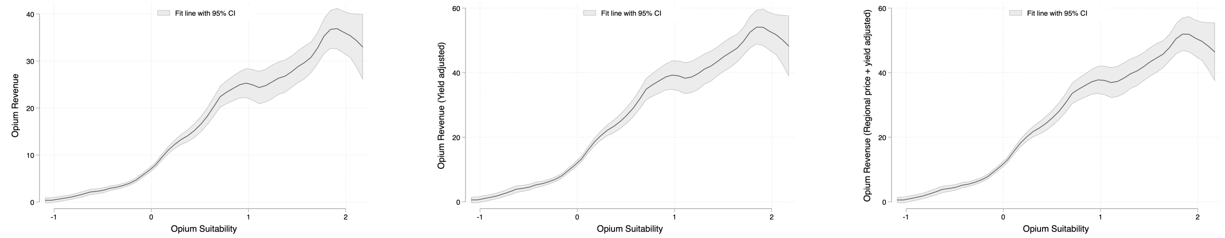
$$viol_{d,t} = \alpha_d + \gamma_t + \delta_{pz}t + \beta_1 endowment_{d,t} + \Lambda X_{d,t} + \epsilon_d, \tag{3}$$

where $viol_{d,t}$ is the level of violence (per capita) for a given district, d, and year, t. These violence levels are calculated in the post-harvest fighting season, following the sequence illustrated in Figure 3. We study a range of violence measures, each following the same benchmark specification. α_d and γ_t capture district and year fixed effects, accounting for district-specific omitted variables that remain fixed over time and time-varying common shocks, including troop surges. $\delta_{pz}t$ denotes prize zone specific time trends. These trends, following the specification in Dube and Vargas (2013), account for potential omitted variables across opium producing regions that vary with time. We also include an array of additional district-specific control variables, denoted by $\Lambda X_{d,t}$. These include, for example, enhanced market access due to expansion of the road network, allocation of small-scale development projects, and the use of coercive tactics by rebels during the planting and growing seasons. Regressions are weighted by population.

The primary quantity of interest in Equation 3 is β_1, which captures the effect of $endowment_{d,t}$ on the violence outcomes. There are several approaches that can be taken to measure $endowment_{d,t}$. The first is to simply study $Opium_{d,t} \times Price_t$. Using this measure would tell us the association between intensity of violence in the fighting season and potential revenue from the opium trade. Although any given district is unlikely to be a price-maker (i.e., $Price_t$ is unlikely to be substantively biased), variation in $Opium_{d,t}$ is only exogenous under the condition that potential sources of bias are captured by the fixed effects and controls. This is unlikely to be the case. Instead, we prefer to capture $endowment_{d,t}$ using $Suitability_{d,t} \times AggProd_{t-1}$. This term captures exogenous variation in suitability driven by climatic and soil conditions as denoted in Equation 1 and standardized as in Equation 2. Suitability is then weighted by time series variation in aggregate production in the prior year, $AggProd_{t-1}$. Once inverted, this captures plausibly exogenous variation in prices due to broader market dynamics: higher aggregate production in the prior year is negatively associated with prices in the current year.

In Figure 4, we illustrate the bivariate fit between various measures of revenue and our preferred suitability measure. Notice that the two measures are strongly positively correlated. In Table A-4 (see page A-18 of the online appendix) we demonstrate that this relationship is robust to the panel model specification introduced in Equation 3. Although these results

Figure 4: Robust Association between Endogenous Revenue Measures and Exogenous Suitability Parameter



(a) Opium revenue

(b) Yield adjusted

(c) Regionally adjusted

Notes: figures depict local polynomial regressions evaluating the relationship between various measures of potential opium revenue and exogenous variation in opium suitability, derived using Equations 1 and 2. For ease of visualization, suitability is trimmed at -1.104 and 2.18. Panel (a) depicts log output in hectares by log of the simple average national price. Panel (b) adjusts (a) using an annual yield calibration weight, allowing us to convert hectares under production to estimation kilograms. Panel (c) allows for regional yield adjustment as well as price zone-by-year price changes.

suggest $Suitability_{d,t} \times AggProd_{t-1}$ is a strong instrument for $Opium_{d,t} \times Price_t$, we prefer a reduced form approach. Focusing on the reduced form enables us to more reliably estimate the marginal effects of $endowment_{d,t}$ in the presence of rebel-led surveillance networks as well as sidestep potential (though largely implausible) concerns that agronomic conditions during the growing season directly influence combat tactics during the fighting season (i.e., a violation of the exclusion restriction).

We next turn to Equation 4, where we investigate the role of intelligence gathering:

$$viol_{d,t} = \alpha_d + \gamma_t + \delta_{pz} t + \beta_1 endowment_{d,t} + \beta_2 endowment_{d,t} \times surveillance_d + \Lambda X_{d,t} + \epsilon_d, \quad (4)$$

where notation follows from Equation 3 and the additional term, $endowment_{d,t} \times surveillance_d$ is the marginal effect of endowment shocks in the presence of the surveillance network. To study this marginal effect, we rely on records gathered by security forces in 2006—the first year of our sample—about rebel-led intelligence gathering (i.e., the areas where they attempted to monitor troop movement and base activities). We leverage cross-sectional variation in this measure to address concerns that the surveillance network adjusts endogenously from year to year, significantly complicating estimation in the absence of a plausible district-specific and time-varying instrument for intelligence gathering. Notice that the base term for $surveillance_d$ is absorbed by α_d (district fixed effects) in Equation 4.

5 Evidence

In this section, we introduce a series of related results. First, we focus on the correlation between rebel capacity and conflict production generally. The primary purpose of these results is to evaluate theoretical claims central to prior related work, most notably Dube and Vargas (2013), which link specific economic shocks to capital-intensive activities. In our setting, we apply this general logic to the most capital-intensive conflict type, direct fire engagements, which typically involve at least several fighters deployed in high risk, line-of-sight attacks. Second, we consider the relationship between exogenous variation of endowments and rebel innovation and attack patterns. This section is most closely aligned with the theoretical model and investigates *how* rebels fight. Third, we evaluate the link between economic shocks and combat effectiveness. In particular, we evaluate whether disruptive and potentially fatal attacks increase with revenue and whether this effectiveness is observed for hardened targets (i.e., Coalition forces). In each section, we also review evidence of whether intelligence gathering is a mechanism that enhances the ability of insurgents to produce more intense, more innovative, and more effective attacks.

5.1 Conflict in levels

We begin by studying the impact of rebel capacity on combat engagements generally. These results are presented in Panel A of Table 1. In column 1, we find that combat activity overall increases significantly. The magnitude of this increase is substantial, with a one standard deviation increase in opium suitability increasing combat activity by .2 standard deviations.[2] In columns 2 through 4, we split apart this combat measure, considering the three primary

---

[2] Summary statistics are presented in each table (for outcomes) and in Table A-1 (see page A-2 of the online appendix).

combat types most frequently observed during the conflict. In columns 2 and 3, we find that direct fire and IED explosions significantly increase with potential revenue from the opium trade (.2 and .25 standard deviation increases respectively). In column 4, on the other hand, we find that production of indirect fire attacks is largely unresponsive to positive shocks to rebel capacity. Of these three combat types, direct fire involves the most significant inputs, requiring both labor (a potentially large number of fighters) and capital (armaments to engage in line-of-sight attacks). The input intensity of roadside bombs varies. One the one hand, emplacing and detonating large-scale bombs can be capital-intensive—involving accumulation of bomb making materials, securing a transport vehicle, locating an ideal target site, and planting the device—as well as labor-intensive, involving one or more skilled bomb makers and a fighter on location to trigger the device. On the other, bombs can be small-scale, involve the use of easily available unexploded ordnance, and remote trigger mechanism (that is, no fighter is present when the device is triggered). We lack the technical information on bomb and weapon fragments recovered from the field, which limits our ability to identify more input intensive attacks. However, we anticipate that on average, the input intensity lays between direct fire and indirect fire, which involves the least risk to fighters. This is due to the remote nature of these attacks, which enable fighters using rocket propelled grenades, mortars, and related weapon types to flee the scene of an attack before counterinsurgents can respond directly. We anticipate, applying the general logic of Dube and Vargas (2013), that these various combat types respond differently to endowment shocks due to their distinct input intensities, with the most capital and labor-intensive attacks increasing most sharply.

In Panel B of Table 1, we introduce results from Equation 4, where we are focused on the marginal effects of rebel capacity in the presence of rebel surveillance. Although the baseline effects are consistently positive (though imprecise with the exception of IED explosions), the marginal effect of exogenous variation in potential revenue had a large, precise, and positive effect on combat activity overall, as well as the production of high risk, input-intensive attacks: direct fire and indirect fire. The magnitude of these increases ranges from .3 to .4

standard deviations (with a standard deviation increase in rebel capacity). As with Panel A, relatively low-risk, indirect fire engagements are largely unresponsive to positive endowment shocks, even in areas where insurgents have previously been engaged in intelligence gathering.

Robustness In the notes for Table 1, we note the additional controls included in the main specification. The main specification accounts for time varying effects of opium price zones and irrigation intensity as well as development assistance and changes in market access. In Table A-5 Panels A and B (see page A-19 of the online appendix), we supplement these covariates with a battery of additional factors that could influence opium productivity and combat activity. All regressors are added to the benchmark specification. These additional covariates include time trends for terrain ruggedness, co-ethnic density, and a historical measure of the Taliban's consolidation of control at the end of 1996, when they initially seized control of Kabul and, with it, the central government. We also incorporate measures of coercive violence and intimidation by insurgents as well as counterinsurgent operations by security forces during the planting and growing seasons, including safe house raids and detentions of suspected fighters and collaborators. Taken together, these results suggest a robust correlation between plausibly exogenous variation in potential opium revenue and the production of violence, especially capital- and labor-intensive attacks.

5.2 Combat Innovation and Sophistication

In this section, we focus on how shocks to rebel capacity influence combat innovation, sophistication, and attack patterns. In Table 2, we focus on four measures of combat innovation, coordination, and complexity. In column 1, we examine the impact of rebel capacity on tactical innovation. As we detail above, coalition forces were instructed to document instances of unexpected tactics and procedures used by insurgents. These technological changes could involve new attack formations and novel weapon systems as well as combat engagement with unique targets. The results in column 1 suggest that innovation is increasing in rebel capacity. The increase is equivalent to .24 standard deviations with each standard deviation increase

Table 1: Impact of Rebel Capacity and Surveillance on Combat Outcomes

	(1) Combat	(2) Direct Fire	(3) IED Explosion	(4) Indirect Fire
Panel A: Baseline Effects				
Opium Suitability	0.450*	0.357†	0.094**	-0.001
	(0.228)	(0.202)	(0.029)	(0.006)
Summary Statistics				
Outcome Mean	0.352	0.223	0.071	0.058
Outcome SD	1.227	0.963	0.224	0.228
Model Statistics				
No. of Observations	3582	3582	3582	3582
No. of Clusters	398	398	398	398
R^2	0.580	0.519	0.627	0.597
Panel B: Heterogeneous Effects				
Opium Suitability	0.079	0.036	0.038**	0.005
	(0.058)	(0.046)	(0.013)	(0.007)
Suit. × Rebel Surveillance	0.744†	0.643†	0.111*	-0.011
	(0.416)	(0.366)	(0.053)	(0.012)
Summary Statistics				
Outcome Mean	0.352	0.223	0.071	0.058
Outcome SD	1.227	0.963	0.224	0.228
Model Parameters				
District Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Price Zone Trends	Yes	Yes	Yes	Yes
Irrigation Intensity Trends	Yes	Yes	Yes	Yes
Dev aid: Ag/Irrigation	Yes	Yes	Yes	Yes
Market Access	Yes	Yes	Yes	Yes
Model Statistics				
No. of Observations	3582	3582	3582	3582
No. of Clusters	398	398	398	398
R^2	0.590	0.533	0.634	0.597

Notes: Outcome of interest varies by column and is indicated in the column heading. The quantity of interest is opium suitability. All regressions include district and year fixed effects as well as controls as specified under model parameters. Heteroskedasticity robust standard errors clustered by district are reported in parentheses. Symbols indicate † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

in suitability. In column 2, we investigate whether the use of deceptive technologies—false and hoax explosive devices—also increases in endowments. Indeed, these results suggest a large increase (.24 SD). In column 3, we evaluate how unit breaches with attacks carried by security force insiders respond to potential revenue shocks. We find that these attacks, which typically require cultivation of an asset already present in the security forces or the deployment of fighters to infiltrate security force units, increase in exogenous endowment variation (.15 SD). In column 4, we consider the association between revenue and attack complexity. We measure target complexity using information about the number of targets involved in a given attack. Complex targets typically include attacks on forces during movement as well as infrastructure. Notice that attacks involving complex targets significantly increase with

Table 2: Impact of Rebel Capacity and Surveillance on Combat Innovation, Coordination, and Complexity

	(1) Tactical Innovation	(2) Deceptive Tech	(3) Unit Breach	(4) Complex Target
Panel A: Baseline Effects				
Opium Suitability	0.004*	0.007**	0.001**	0.043**
	(0.002)	(0.002)	(0.000)	(0.016)
SUMMARY STATISTICS				
Outcome Mean	0.003	0.004	0.001	0.014
Outcome SD	0.014	0.021	0.004	0.085
MODEL STATISTICS				
No. of Observations	3582	3582	3582	3582
No. of Clusters	398	398	398	398
R^2	0.262	0.355	0.200	0.385
Panel B: Heterogeneous Effects				
Opium Suitability	-0.001	0.003†	0.000†	0.015*
	(0.001)	(0.002)	(0.000)	(0.007)
Suit. × Rebel Surveillance	0.010*	0.007*	0.001†	0.056*
	(0.004)	(0.003)	(0.001)	(0.026)
SUMMARY STATISTICS				
Outcome Mean	0.003	0.004	0.001	0.014
Outcome SD	0.014	0.021	0.004	0.085
MODEL PARAMETERS				
District Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Price Zone Trends	Yes	Yes	Yes	Yes
Irrigation Intensity Trends	Yes	Yes	Yes	Yes
Dev aid: Ag/Irrigation	Yes	Yes	Yes	Yes
Market Access	Yes	Yes	Yes	Yes
MODEL STATISTICS				
No. of Observations	3582	3582	3582	3582
No. of Clusters	398	398	398	398
R^2	0.276	0.358	0.203	0.398

Notes: Outcome of interest varies by column and is indicated in the column heading. The quantity of interest is opium suitability. All regressions include district and year fixed effects as well as controls as specified under model parameters. Rounded coefficient in column 3, Panel B is .0004. Heteroskedasticity robust standard errors clustered by district are reported in parentheses. Symbols indicate † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

exogenous access to potential revenue (.3 SD).

In Panel B of Table 2, we return to the mechanism at the core of our theoretical model: intelligence gathering. While these measures of innovation, coordination, and complexity generally increase in revenue, the marginal effects of revenue in the presence of rebel-led surveillance are large, positive, and consistently precise. The magnitude of these effects is also substantial, ranging from .2 to .4 standard deviations. This result suggests that the ability to convert fighting capacity into battlefield innovations largely hinges on the presence of networks that enable rebels to gather information about the vulnerabilities of rival forces.

In Table 3, we take these results a step further. In this set of results we consider another implication of the theoretical model, that attack patterns may become more clustered in time or space as rebels gather more information about their opponent's defensive or offensive weaknesses. To study clustering, we consider two novel measurement strategies.

We begin by focusing on temporal clustering—the concentration of attacks in certain windows of time during a given day. We quantify patterns of combat operations using randomization inference and the bootstrap Kolmogorov-Smirnov method developed by Abadie (2002). Using this approach, lower values of the dependent variable reveal attack patterns that are more easily differentiated from randomness; i.e., they are more clustered. We provide a more detailed description of the method in the online appendix (see page A-13). In columns 1 and 2 of Panel A, we consider whether combat clustering increases in revenue. For interpretation, recall that lower values indicate more robust clustering patterns. Therefore, we anticipate a negative coefficient on our measure of resource endowments. Columns 1 and 3 do not include unit fixed effects since these measures of clustering yield an unbalanced panel. In columns 2 and 4, we add unit fixed effects to the model specification. Indeed, though we do find a negative association with respect to temporal clustering, it is imprecise in Panel A. In Panel B, however, we find evidence of a large, and precisely estimated increase in clustering with respect to endowments in areas where rebels may acquire information through base and troop surveillance (.2 SD increase).

In Panel A, columns 3 and 4, we turn to another measure of attack clustering. In these specifications, we are focused on spatial clustering. To investigate spatial clustering, we develop a 5 kilometer × 5 kilometer grid of Afghanistan, linking observed combat activity to this grid. We then calculate an *Index of Dispersion* for each district-fighting season by linking grid cells to their corresponding parent administrative unit (Perry and Mead, 1979). This type of index is common in the study of spatial point processes (e.g., seedling dispersion). Higher values indicate that the spatial pattern was highly unlikely to have occurred by random chance and exhibits characteristics of uneven density (i.e., spatial clustering of attacks). Given the flipped interpretation of the outcome variable (relative temporal clustering), we expect a positive correlation with suitability if our model extends to spatial allocation of attacks as well as temporal clustering. Notice, in columns 3 and 4, spatial clustering of attacks increases with potential revenue (.27 SD). In Panel B, we consider whether these revenue effects are largest in areas where rebels have been able to conduct surveillance historically. Indeed, we find that the margin effect of revenue in these areas is large and positive, indicating that clustering increases the most with exogenous endowments when rebels are capable of coordinated intelligence gathering (.48 SD).

Robustness In Panels A and B of Tables A-6 and A-7 (see pages A-20 and A-21 of the online appendix, respectively), we supplement the main specification with additional covariates, including trends for terrain ruggedness, coethnicity, and historical Taliban control as well as measures of insurgent coercion and counterinsurgent operations for thwarting insurgent coordination and surveillance. The results overall are highly consistent, with innovation, deception, complexity, and clustering sharply increasing in suitability overall and heterogeneously with rebel-led surveillance activity. Although infiltration is precisely correlated with our measure of potential revenue overall, the heterogeneous effect with respect to surveillance loses precision at the 10 percent threshold ($t = 1.54$).

5.3 Combat Losses and Casualties

In this section, we turn to a first-order question regarding civil conflict: do resource endowments increase the effectiveness of rebel attacks? We investigate this question in Table 4, Panel A. In column 1, we focus on variation in combat events that caused vehicular damage or security forces casualties. In column 2, we focus more narrowly on events involving casualties. In columns 3 and 4, we focus on damage and casualties inflicted on harder targets, specifically Coalition forces. Notice that all four measures of combat losses and casualties are increasing in revenue. Also importantly, the coefficient magnitudes are larger when we focus on hard targets (columns 3 and 4), indicating that combat effectiveness is increasing against otherwise more battle ready units. Comparing columns 1 and 3, the magnitude of these differences is substantial: a .14 standard deviation increase versus a .29 standard de-

Table 3: Impact of Rebel Capacity and Surveillance on Attack Clustering

	(1) Temporal	(2) Temporal (TWFE)	(3) Spatial	(4) Spatial (TWFE)
Panel A: Baseline Effects				
Opium Suitability	-0.659	-0.349	496.815*	346.963†
	(0.448)	(0.510)	(236.005)	(198.850)
SUMMARY STATISTICS				
Outcome Mean	-7.048	-7.110	343.140	350.022
Outcome SD	4.355	4.368	1072.541	1083.435
MODEL STATISTICS				
No. of Observations	1467	1435	1467	1435
No. of Clusters	266	234	266	234
R^2	0.116	0.387	0.186	0.599
Panel B: Heterogeneous Effects				
Opium Suitability	0.093	0.471	80.005	37.973
	(0.246)	(0.361)	(68.404)	(67.666)
Suit. × Rebel Surveillance	-1.558*	-1.529†	867.764*	576.607†
	(0.728)	(0.834)	(369.321)	(315.301)
SUMMARY STATISTICS				
Outcome Mean	-7.048	-7.110	343.140	350.022
Outcome SD	4.355	4.368	1072.541	1083.435
MODEL PARAMETERS				
District Fixed Effects	No	Yes	No	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Price Zone Trends	Yes	Yes	Yes	Yes
Irrigation Intensity Trends	Yes	Yes	Yes	Yes
Dev aid: Ag/Irrigation	Yes	Yes	Yes	Yes
Market Access	Yes	Yes	Yes	Yes
MODEL STATISTICS				
No. of Observations	1467	1435	1467	1435
No. of Clusters	266	234	266	234
R^2	0.144	0.391	0.259	0.609

Notes: Outcome of interest varies by column and is indicated in the column heading. The quantity of interest is opium suitability. All regressions include year fixed effects as well as controls as specified under model parameters. Even columns include district fixed effects. Heteroskedasticity robust standard errors clustered by district are reported in parentheses. Symbols indicate † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

viation increase for the equivalent shock to suitability. These differences are even larger in standardized terms regarding columns 2 and 4, where size of the effect for coalition targets is roughly three times larger than government targets generally (.09 SD versus .28 SD). Panel B, where we introduce marginal effects for intelligence gathering, suggests the increased combat impact is greatest in areas where rebels have previously engaged in surveillance. In Table A-8 Panels A and B (see page A-22 of the online appendix), we present results from a more saturated set of model specifications. The additional covariates follow the discussion above, with the supplemental models accounting for time trends with respect to terrain ruggedness, coethnicity, and historical control by the Taliban. These models also account for the possibility that coercion by insurgents and efforts to disrupt rebel operations by

Table 4: Impact of Rebel Capacity and Surveillance on Combat Losses and Casualties

	(1) Disrupt: Govt	(2) Casualties: Govt	(3) Disrupt: Coal	(4) Casualties: Coal
Panel A: Baseline Effects				
Opium Suitability	0.061*	0.029†	0.071**	0.039*
	(0.024)	(0.016)	(0.026)	(0.016)
SUMMARY STATISTICS				
Outcome Mean	0.079	0.060	0.031	0.017
Outcome SD	0.255	0.202	0.142	0.083
MODEL STATISTICS				
No. of Observations	3582	3582	3582	3582
No. of Clusters	398	398	398	398
R^2	0.615	0.598	0.449	0.436
Panel B: Heterogeneous Effects				
Opium Suitability	0.012	-0.006	0.025*	0.009
	(0.010)	(0.008)	(0.010)	(0.006)
Suit. × Rebel Surveillance	0.099*	0.070**	0.092†	0.061*
	(0.042)	(0.027)	(0.048)	(0.029)
SUMMARY STATISTICS				
Outcome Mean	0.079	0.060	0.031	0.017
Outcome SD	0.255	0.202	0.142	0.083
MODEL PARAMETERS				
District Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Price Zone Trends	Yes	Yes	Yes	Yes
Irrigation Intensity Trends	Yes	Yes	Yes	Yes
Dev aid: Ag/Irrigation	Yes	Yes	Yes	Yes
Market Access	Yes	Yes	Yes	Yes
MODEL STATISTICS				
No. of Observations	3582.000	3582.000	3582.000	3582.000
No. of Clusters	398.000	398.000	398.000	398.000
R^2	0.619	0.601	0.461	0.452

Notes: Outcome of interest varies by column and is indicated in the column heading. The quantity of interest is opium suitability. All regressions include district and year fixed effects as well as controls as specified under model parameters. Heteroskedasticity robust standard errors clustered by district are reported in parentheses. Symbols indicate † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

government forces are correlated with opium production and downstream combat tactics. Overall, these results suggest that damage and casualties to government forces overall, and hardened coalition targets specifically, increase with exogenous variation in opium suitability and surveillance by rebels.

6 Conclusion

Rebel tactics are an overlooked feature of internal warfare. Our paper aims to address this gap, coupling insights about the economic effects of resource endowments during conflict with theories of information dynamics. The argument we advance is that resource endowments, especially when coupled with information about potential enemy vulnerabilities, enhance the ability of insurgents to initiate violence, engage in tactical and technological innovation, and improve the efficiency and effectiveness of their attacks. Our model of these dynamics extends prior work on Colonel Blotto-type games and has implications for variety of important conflict dynamics, including when and where combatants attack and the complexity of their attack patterns. Our focus on the role of intelligence gathering by non-state actors develops observations in other qualitative and ethnographic work, including Kalyvas (2006), and deepens the set of theoretical models of insurgent behavior during conflict by highlighting how information flow—to rebels—may undermine government operations.

Studying these dynamics requires unusually rich microdata. Using data collected during Operation Enduring Freedom in Afghanistan, we study a range of combat outcomes, including previously unreleased measures of tactical innovation, deceptive weaponry, and target complexity. We couple these outcomes with plausibly exogenous, microlevel measures of opium suitability and leverage the industrial organization of the Taliban, including their highly institutionalized taxation system, to estimate the impact of potential revenue from the drug trade on combat tactics in the subsequent fighting season. Our data also yield a unique opportunity to assess the underlying mechanism suggested by our theoretical model: intelligence gathering.

Consistent with the argument, we find that as rebels accumulate fighting capacity, their attacks increase in intensity, become more sophisticated, and yield more government combat losses and casualties. Overall, we find that these battlefield consequences of resource endowments are greatest in areas where rebels have engaged in troop and base surveillance operations. While prior quantitative and theoretical work has detailed the role of information in shaping counterinsurgency effectiveness (i.e., civilian tips to government forces), our model and results help clarify the value and consequences of rebel-led information gathering through surveillance.

In the aftermath of the Taliban's military offensive and capture of Kabul in 2021, it is important to reflect on the broader lessons that can be drawn from this study. As Barnett

Rubin, a former State Department adviser on Afghanistan has noted, the narcotics sector is "the country's largest industry except for war." Afghanistan's economic development after the war industry subsides will likely hinge both on how the Taliban manage their relations with foreign aid donors and whether a government under Taliban rule promotes or prohibits opium production. Mullah Omar's struggle with this diplomatic and economic trade-off is likely to reflect a future policy dilemma: substantive efforts to appease international counter-narcotics demands will undermine rural economies and weaken the Taliban's hold on power. The very resource endowment that enabled their survival after removal from power and rise to power after the U.S. withdrawal may constrain their ability to rule.

Our findings are more broadly relevant to how civil conflicts are fought and how they end. We present evidence highlighting the role of these resource endowments in shaping how the group engaged in violence, developed innovative, sophisticated combat techniques, and waged combat effectively against technologically superior coalition forces, echoing results in Fetzer et al. (2021) about the Taliban's strategic use of violence during the phased withdrawal of international troops. From a more global perspective, our paper suggests that illicit economies can enable armed groups to survive and outlast even high-capacity international military forces. In this sense, the Afghan experience mirrors other civil conflicts, including Colombia, Iraq, and the Philippines, where rebels developed sophisticated institutions to gather or capture resources from civilians and private firms, monitor state forces, and engage in political violence.

References

Abadie, Alberto. 2002. "Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models." *Journal of the American Statistical Association* 97(457):284–292.

Bazzi, Samuel and Christopher Blattman. 2014. "Economic Shocks and Conflict: Evidence from Commodity Prices." *American Economic Journal: Macroeconomics* 6(4):1–38.

Berman, Eli and Aila M. Matanock. 2015. "The Empiricists' Insurgency." *Annual Review of Political Science* 18:443–64.

Berman, Nicolas, Mathieu Couttenier, Dominic Rohner and Mathias Thoenig. 2017. "This Mine Is Mine! How Minerals Fuel Conflicts in Africa." *American Economic Review* 107(6):1564–1610.

Besley, Timothy and Torsten Persson. 2011. "The Logic of Political Violence." *The Quarterly Journal of Economics* 126(3):1411–1445.

Blackett, D. W. 1958. "Pure Strategy Solutions of Blotto Games." *Naval Research Logistics Quarterly* 5(2):107–109.

Condra, Luke N and Jacob N Shapiro. 2012*a*. "Who takes the blame? The strategic effects of collateral damage." *American journal of political science* 56(1):167–187.

Condra, Luke N. and Jacob N. Shapiro. 2012*b*. "Who Takes the Blame? The Strategic Effects of Collateral Damage." *American Journal of Political Science* 56(1):167–187.

Condra, Luke N., James D. Long, Andrew C. Shaver and Austin L. Wright. 2018. "The Logic of Insurgent Electoral Violence." *American Economic Review* 108(11):3199–3231.

Dell, Melissa, Benjamin Jones and Benjamin Olken. 2014. "What do we learn from the weather? The new climate-economy literature." *Journal of Economic Literature* 52(3):740–798.

Dell, Melissa and Pablo Querubin. 2018. "Nation Building Through Foreign Intervention: Evidence from Discontinuities in Military Strategies." *Quarterly Journal of Economics* 133(2):701–764.

Dube, Oeindrilla and Juan Vargas. 2013. "Commodity Price Shocks and Civil Conflict: Evidence from Colombia." *Review of Economic Studies* 80(4):1384–1421.

Farrell, Graham and John Thorne. 2005. "Where have all the flowers gone?: Evaluation of the Taliban crackdown against opium poppy cultivation in Afghanistan." *International Journal of Drug Policy* 16(2):81–91.

Felbab-Brown, Vanda. 2006. "Kicking the opium habit?: Afghanistan's drug economy and politics since the 1980s: Analysis." *Conflict, Security & Development* 6(2):127–149.

Fetzer, Thiemo, Pedro C. L. Souza, Oliver Vanden Eynde and Austin L. Wright. 2021. "Security Transitions." *American Economic Review* 111(7):2275–2308.

Giustozzi, Antonio. 2019. *The Taliban at War: 2001-2018.* Oxford University Press.

Kalyvas, Stathis N. 2006. *The Logic of Violence in Civil War.* New York: Cambridge University Press.

Konig, Michael D., Dominic Rohner, Mathias Thoenig and Fabrizio Zilibotti. 2017. "Networks in Conflict: Theory and Evidence From the Great War of Africa." *Econometrica* 85(4):1093–1132.

Kovenock, Dan and Brian Roberson. 2012. "Conflicts with Multiple Battlefields," in Michelle R. Garfinkel and Stergios Skaperdas (eds.). Oxford Handbook of the Economics of Peace and Conflict, Oxford University Press, New York. pp. 503-531.".

Limodio, Nicola. 2019. "Terrorism Financing, Recruitment and Attacks: Evidence from a Natural Experiment." *Chicago Booth Research Paper No. 32* .

Mejía, Daniel and Pascual Restrepo. 2014. "Bushes and Bullets: Illegal Cocaine Markets and Violence in Colombia." Working paper.

Na, Neo Wee. 2018. "Opium Production and Countering Terrorism Financing in Afghanistan: Lessons from Thailand's Royal Projects." *Counter Terrorist Trends and Analyses* 10(2):1–5.

Perry, J. N. and R. Mead. 1979. "On the Power of the Index of Dispersion Test to Detect Spatial Pattern." *Biometrics* 35(3):613–622.

Peters, Gretchen S. 2009. *The Taliban and the Opium Trade.* New York: Columbia University Press pp. 7–22.

Powell, Robert. 2007. "Defending against Terrorist Attacks with Limited Resources." *American Political Science Review* 101(3):527–541.

Powell, Robert. 2013. "Monopolizing Violence and Consolidating Power." *Quarterly Journal of Economics* 128(2):807–859.

Saha, Suranjana, Shrinivas Moorthi, Hua-Lu Pan, Xingren Wu, Jiande Wang and et al. 2010. "The NCEP Climate Forecast System Reanalysis." *Bulletin of the American Meteorological Society* 91(8):1015–1057.

Sexton, Renard. 2016. "Aid as a Tool against Insurgency: Evidence from Contested and Controlled Territory in Afghanistan." *American Political Science Review* 110(4):731–749.

Shaver, Andrew and Austin Wright. 2016. "Are Modern Insurgencies Predictable? New. Evidence from the Afghanistan and Iraq Wars." Working paper.

Vanden Eynde, Oliver. 2018. "Targets of Violence: Evidence from India's Naxalite Conflict." *The Economic Journal* 128(609):887–916.

Weidmann, Nils B. 2016. "A Closer Look at Reporting Bias in Conflict Event Data." *American Journal of Political Science* 60(1):206–218.

Weinstein, Jeremy M. 2007. *Inside Rebellion: The Politics of Insurgent Violence*. Cambridge: Cambridge University Press.

Wood, Reed M. 2014. "Opportunities to Kill or Incentives for Restraint? Rebel Capabilities, the Origins of Support, and Civilian Victimization in Civil War." *Conflict Management and Peace Science* 31(5):461–480.

Appendix

Contents

Appendix A – Summary Statistics

Table A-1: Summary Statistics

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Panel A: Combat outcomes				
Combat, pc	0.352	1.227	0	33.183
Direct fire, pc	0.223	0.963	0	29.436
IED explosions, pc	0.071	0.224	0	4
Indirect fire, pc	0.058	0.228	0	7.485
Panel B: Innovation outcomes				
Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Tactical innovation, pc	0.003	0.014	0	0.466
Deceptive weapons technology, pc	0.004	0.021	0	0.447
Unit infilitration, pc	0.001	0.004	0	0.189
Complex target, pc	0.014	0.085	0	1.287
Panel C: Battlefield effectiveness outcomes				
Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Government disruption, pc	0.079	0.255	0	6.231
Government casualties, pc	0.06	0.202	0	5.594
Coalition disruption, pc	0.031	0.142	0	2.991
Coalition casualties, pc	0.017	0.083	0	1.84
N	3582			
Panel D: Innovation outcomes (unbalanced panel)				
Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Temporal clustering, combat	-7.048	4.355	-42.951	0
Spatial clustering, combat	343.14	1072.541	0	18613.631
N	1467			
Panel E: Measures of rebel capacity				
Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Opium suitability	-0.03	0.606	-2.151	3.608
Opium revenue	8.256	14.041	0	55.562
Opium revenue, yield adjusted	13.567	21.803	0	75.518
Opium revenue, regional price/yield adj	13.194	21.208	0	73.587
N	3582			

Notes: summary statistics are calculated for the sample studied in the main estimating equation. Per capita outcomes are measured per 1000 residents using population data collected by Afghanistan's Central Statistics Organization

Appendix B – Theory

Technology of conflict has long been an active area of theoretical modelling (Kress, 2012). Optimal allocation of attacking and defensive resources has been studied in the Colonel Blotto-type games (Borel, 1921; Blackett, 1958; Powell, 2007). See Golman and Page (2009) and Konrad and Kovenock (2009) for recent advances and Kovenock and Roberson (2012) for an excellent survey.

Proof of Proposition 1

For event H, the complement is denoted H'. We introduce auxiliary random variables T_i, S_i taking two values 0 and 1; $T_i = 1$ means that i-th target is defended (i.e., $r_i = 1$), $T_i = 0$ means that there is no defense at time i (i.e., $r_i = 0$). (For technical reasons, we need this subtle distinction between a government strategy $(r_1, ..., r_n)$ and random variables T_i.) Similarly, $S_i = 1$ means that the test of target i shows it as defended (i.e., $s_i = 1$), and $S_i = 0$ means that the target i tests as vulnerable (i.e., $s_i = 0$). (Again, the difference is between signal values and the random variable that takes these values.) In the absence of index i, S and T correspond to any target.

Recall that $P(S|T) = P(S'|T') = \theta$.[1] Finally, let C denote the event that attack is successful during a target with an attack. We assume that $P(C|T') = p$ and $P(C|T) = 0$.

We will show that for rebels the optimal strategy looks as follows. If $a < x$, the attacks are allocated uniformly at random among the vulnerable targets. If $a > x$ then there is a threshold value, a function of parameters of the model and x, that determines how many should be allocated to the targets; above that, they start to put into the targets that tested "defended", again uniformly.

First, observe that cases $x = 0$ and $x = n$ are trivial: there is no information to infer, so the optimal strategy for rebels is to allocate attacks uniformly across n targets.

After n targets are tested, a (vector) signal $s = (s_1, .., s_n)$ is obtained and value $N = x$, the number of targets that tested vulnerable is produced. In other words, a (random) partition of a set J into two sets, $J^-(x)$ - targets tested vulnerable, and its complement $J^+(x)$, is obtained. Define $N_1 = |J^-(x) \cap L|$, the number of defended sites that tested vulnerable, i.e., the number of *"false positives"*; and $N_2 = |J^-(x) \cap U|$, the number of correct vulnerable signals. The total number of sites that tested vulnerable is $N \equiv N_{n,r} = N_1 + N_2$.

N_1 and N_2 are independent binomial random variables. Denote $p_i(j)$ the p.m.f. of $N_i, i = 1, 2$, and $p(j|l, p), j = 0, 1, ..., l$ – the p.m.f. of a binomial random variable with l trials and probability of success p. Then $p_1(j) = p(j|r, 1 - \theta)$ and $p_2(j) = p(j|n - r, \theta)$.

Now the p.m.f. of N, $h_{n,r}(x)$, $0 \leq x \leq n$, can be calculated by the standard discrete convolution formula.

---

[1] All the results go through with the arbitrary parameters $\alpha = P(S|T)$, $\beta = P(S'|T')$ subject to $\alpha + \beta > 1$, i.e., that the test is informative. A standard interpretation for α and β as $\alpha = P(\text{positive test}|\text{disease}) =$ sensitivity; and $\beta = P(\text{negative test}|\text{no disease})=$ specificity, two important characteristics of any statistical test.

$$h_{n,r}(x) = \sum_{0 \le j \le r,\ 0 \le x-j \le n-r} p_1(j) p_2(x-j).$$

We start by observing that $P(C_i = 1|T_i = 1) = 0$ and $P(C_i = 1|T_i = 0, a_i) = p(a_i)$, where a_i is the number of attacks launched against target i, and $p(a)$ is the success function, the probability of at least one successful attack on a target which faces u attacks.

As we assumed that the success is independent across attacks, $p(a) = 1 - (1-p)^a$. The function $p(a)$ is increasing and upward concave, and the function $\Delta p(a) \equiv p(a+1) - p(a)$ is decreasing. The diminishing effect of each extra attack will play an important role in determing the optimal strategy.

We start with a lemma.

Lemma A1 *The posterior probability of signal distribution is uniform conditional on the number of signals "vulnerable" x :*

$$P(s_1, ..., s_n) = P(N = x) / \binom{n}{x}. \tag{A1}$$

The posterior probability that target i is vulnerable conditional on the full vector signal $(s_1, ..., s_n)$ is equal to the conditional probability that target i is vulnerable conditional only on the individual signal s_i and the total number of "vulnerable" signals x.

$$P(T_i = 0|s_1, ..., s_n) = P(T_i = 0|s_i, N = x). \tag{A2}$$

Proof. (A1) is straightforward. The left side of formula (A2) can be written as

$$P(T_i = 0) P(s|T_i = 0)/P(s) = P(s_i|T_i = 0) P(s_{-i}|T_i = 0)/P(s),$$

where s_{-i} is vector s without coordinate s_i. Using (A1), we can replace $P(s) = P(N = x)/\binom{n}{x}$. The right side of formula (A2) can be written as

$$\frac{P(T_i = 0, s_i, N = x)}{P(s_i, N = x)} = \frac{P(T_i = 0) P(s_i|T_i = 0) P(N = x|s_i, T_i = 0)}{P(N = x) P(s_i|N = x)}.$$

Let $s_i = 0$. Then, on the left-hand side, using (A1) for a problem with $n-1$ targets and k attacks, we have

$$P(s_{-i}|T_i = 0) = P_{n-1,k}(N = x-1) / \binom{n-1}{x-1}.$$

In the right-hand side we have $P(s_i|N = x) = x/n$ and

$$P(N = x|s_i = 0, T_i = 0) = P_{n-1,k}(N = x-1).$$

Finally, since $\binom{n}{x} = \binom{n-1}{x-1} n/x$, we obtain that after all reductions the left and the right sides of formula (A2) coincide. The proof for the case $s_i = 1$ is similar. ■

The formula (A2) is at the heart of the intuition behind our main results. The optimal strategy $\pi(\cdot|x)$ of the attacker depends on function $p(a)$ and on the critical ratio

$$\rho_{n,r}(x|\theta) \equiv \frac{p^-(x)}{p^+(x)}.$$

The ratio reflects the relative vulnerability of targets with $s = 0$ and $s = 1$, which, in turn, depends on parameters n, r and θ. Then, the ratio determines the threshold $\overline{a}(x)$ for each x.

Our next goal is to establish that $\rho_{n,r}(x) \geq 1$.

Lemma A2 *(a) The probabilities* $p^-(x) \equiv P(T'|S', x)$ *and* $p^+(x) \equiv P(T'|S, x)$ *for* $0 < x < n$ *are given by formulas*

$$\begin{aligned} p^-(x) &= \frac{r}{x} * \theta * \frac{h_{n-1,r}(x-1)}{h_{n,r}(x)}, \\ p^+(x) &= \frac{r}{n-x} * (1-\theta) * \frac{h_{n-1,r}(x)}{h_{n,r}(x)}. \end{aligned}$$

(b) The ratio $\rho_{n,r}(x|\theta)$, $0 < x < n$, *is given by the formula*

$$\rho_{n,r}(x|\theta) = \frac{\theta}{1-\theta} \frac{n-x}{x} \frac{h_{n-1,r}(x-1)}{h_{n-1,r}(x)}. \tag{A3}$$

Proof. (a) For the sake of brevity, we denote $P(\cdot|N = x) \equiv P(\cdot|x)$ and event $D(x) = (N = x) = D$. By definition $p^-(x) \equiv P(T'|S', x) = P(T'S', x)/P(S', x)$, and $p^+(x) \equiv P(T'|S, x) = P(T'S, x)/P(S, x)$. We have

$$\begin{aligned} P(T'S'D) &= P(T')P(S'D|T') = P(T')P(S'|T')P(D|T'S'), \\ P(T'SD) &= P(T')P(SD|T') = P(T')P(S|T')P(D|T'S). \end{aligned} \tag{A4}$$

We also have $P(T') = \frac{r}{n}$, $P(S'|T') = \theta$, and $P(S|T') = 1 - \theta$. Therefore, to prove the result, it is sufficient to show that $P(D|T'S') = h_{n-1,r}(x-1)$, $P(D|T'S) = h_{n-1,r}(x)$, $P(S'D) = h_{n,r}(x) * \frac{x}{n}$, and $P(SD) = h_{n,r}(x) * \frac{n-x}{n}$.

To show that $P(D|T'S') = h_{n-1,r}(x-1)$, observe that if $N = x$, and a particular target has no defense, T', and produced vulnerable signal, S', then in the remaining $n-1$ targets there are r defenses with $x-1$ vulnerable signals. Similarly, to show that $P(D|T'S) = h_{n-1,r}(x)$, observe that if $N = x$, and a particular target has no defense, T', and produced "defended" signal, S, then in the remaining $n-1$ targets there are r defenses with x vulnerable signals. To demonstrate that $P(S'D) = h_{n,r}(x) * \frac{x}{n}$, note that $P(S'D) = P(D)P(S'|D)$, $P(D) = h_{n,r}(x)$ and $P(S'|D) = \frac{x}{n}$, the probability for one vulnerable signal among x to be in a particular target. Similarly, $P(SD) = P(D)P(S|D)$, and $P(S|D) = \frac{n-x}{n}$, the probability for one defended signal among $n-x$ to be in a particular target.

(b) is a straightforward corollary to (a). ■

Lemma A3 *(a)* $\rho(x) > 1$.

(b) Functions $\rho_{n,n-1}(x) = \frac{\theta^2}{1-\theta^2}$ *for all* x*, while functions* $\rho_{n,r}(x|\theta)$ *for* $r < n-1$ *are monotonically increasing in* x *for* $0 < x < n$*.*

(c) Functions $\rho_{n,r}(x|\theta)$ *are monotonically decreasing for all fixed* r*,* $0 < x < n$ *when* n *is increasing.*

Proof. (a) Let $D_{n,r}$ be a sum of n Bernoulli random variables, r of which have parameter $1-\theta$, and $n-r$ of which have parameter $\theta > 1-\theta$, $0 \le r \le n$. Let

$$\begin{aligned} h_{n,r}(x) &= P(D_{n,r} = x),\ 0 \le x \le n; \\ f_{n,r}(x) &= \frac{h_{n,r}(x-1)}{h_{n,r}(x)},\ 1 \le x \le n; \end{aligned}$$

$$\rho_{n,r}(x) = \frac{n+1-x}{x}\frac{\theta}{1-\theta}f_{n,r}(x), 1 \le x \le n.$$

We assume also that $f_{n,r}(0) = \rho_{n,r}(0) = 0$.

Denote $c = \frac{\theta^2}{1-\theta^2}$. Then $\theta > \frac{1}{2}$ implies $c > 1$.

It is easy to check that

$$\begin{aligned} \rho_{n,0}(x) &= 1,\ \text{for } 1 \le x \le n\,; \\ \rho_{n,n}(x) &= c,\ \text{for } 1 \le x \le n\,; \\ \rho_{n,r}(n) &= \frac{rc+n-r}{n} = 1 + \frac{r}{n}(c-1) \text{ for } 0 \le r \le n\,. \end{aligned} \tag{A5}$$

Using the total probability formula, we obtain a recursive relationship:

$$f_{n,r}(x) = \frac{1-\theta+\theta f_{n-1,r}(x-1)}{\frac{1-\theta}{f_{n-1,r}(x)} + \theta} \text{ for } 1 \le r,\ x \le n-1\,. \tag{A6}$$

This implies that

$$\rho_{n,r}(x) = \frac{n-x+1+(x-1)\rho_{n-1,r}(x-1)}{\frac{n-x}{\rho_{n-1,r}(x)} + x} \text{ for } 1 \le r,\ x \le n-1. \tag{A7}$$

(A5) is the induction base. By induction, the numerator in (A7) is strictly increasing, and the denominator is strictly decreasing, and hence the right side in (A7) is strictly increasing and depends only on c. When c grows from 1 to ∞ (that is, θ is increasing from 0 to $\frac{1}{2}$), it is strictly increasing from 1.

(b) Now we can show that for fixed $n \ge 2,\ 1 \le r \le n-1,\ c > 1$, function $\rho_{n,r}(x)$ is strictly increasing in x.

Again, we use induction by n. Let $\rho(x) = \rho_{n-1,r}(x),\ H(x) = n - x + x\rho(x)$. Then, by (A7),

$$\rho_{n,r}(x) = \rho(x)\frac{H(x-1)}{H(x)}, \quad \rho_{n,r}(x+1) - \rho_{n,r}(x) = \frac{C(x)}{H(x+1)H(x)}\,,$$

where

$$C(x) \;\; = \;\; \rho(x+1)\left[(n-x)^2 + 2x(n-x)\rho(x) + x^2\rho^2(x)\right].$$

We can check, using (A6) and (A7), that

$$\rho_{2,1}(2) = \frac{1+c}{2} > \rho_{2,1}(1) = \frac{2c}{1+c},$$

$$\rho_{3,1}(3) = \frac{2+c}{3} > \rho_{3,1}(2) = \frac{1+2c}{2+c} > \rho_{3,1}(1) = \frac{3c}{1+2c},$$

$$\rho_{3,2}(3) = \frac{1+2c}{3} > \rho_{3,2}(2) = c\frac{2+c}{1+2c} > \rho_{3,2}(1) = \frac{3c}{2+c},$$

and hence the proof is complete for $n = 2$ and $n = 3$.

For $n \geq 4$ and $2 \leq x \leq n-2$, we have $x(n-x) - n \geq 0$, and by induction, it follows that

$$C(x) \;\; > \;\; \rho(x)(\rho(x+1)-1)(\rho(x)-1) > 0$$

for $2 \leq x \leq n-2$.

To prove (c), it remains to show that $\rho_{n-r,r}(2) - \rho_{n-r,r}(1) > 0$ and $\rho_{n-r,r}(n) - \rho_{n-r,r}(n-1) > 0$. Using the total probability formula,

$$\rho_{n,r}(n-1) \;\; = \;\; \frac{1}{n-1}\,\frac{(n-r)(n-r-1) + 2(n-r)rc + r(r-1)c^2}{n-r+cr}. \tag{A8}$$

Then, using (A6), we obtain

$$\rho_{n,r}(n) - \rho_{n,r}(n-1) = \frac{(n-r)r(c-1)^2}{n(n-1)(n-r+cr)} > 0\,.$$

Similarly,

$$\rho_{n,r}(2) - \rho_{n,r}(1) = \frac{(n-r)r(c-1)^2}{[c^2(n-r)(n-r-1) + 2(n-r)rc + r(r-1)](r+c(1-r))}\,.$$

■

Let $B^-(s) = \{i : s_i = 0\}$ and $B^+(s) = \{i : s_i = 1\}$. Then, using (A2), we obtain that, given a strategy $\pi = (a_1, ..., a_n)$ and any signal s with $N(s) = x$, the expected value of a strategy π is

$$w(\pi|x) = 1 - \Pi_{j=1}^{n}(1 - P(C_j|a_j, s_j, x)).$$

Let $U^- \equiv U^-(\pi|s) = \{a_j, j \in B^-(s)\}$ and $U^+ \equiv U^+(\pi|s) = \{a_j \in B^+(s)\}$ be two possible sets of the values of a_j at vulnerable and non-vulnerable targets. Formula (A2) immediately implies that all strategies obtained by permutations of sets (U^-, U^+) among corresponding targets have the same value.

We prove later that the upward concavity of function $p(a)$ implies that the optimal strategy has the property that the number of attacks against any pair of targets with the same signal is the same or almost the same: as the number of attacks is integer, there might be a difference of one attack between two targets with the same signals. Let us assume for simplicity that both equalities hold: $a_i = u^-, i \in B^-(s)$ and $a_i = u^+, i \in B^+(s)$.

Let us do the following transformation:

$$\ln(1 - w(\pi|x)) \;=\; n - p^-(x) \sum_{i \in B^-(s)} p(a_i) - p^+(x) \sum_{i \in B^+(s)} p(a_i).$$

Now, the same strategy π that maximizes the function $v(\pi|x) = p^-(x) \sum_{i \in B^-(s)} p(a_i) - p^+(x) \sum_{i \in B^+(s)} p(a_i)$ is maximizing the strategy value $\ln(1 - w(\pi|x))$.

The following lemma concludes the proof of Proposition 1. The optimal strategy has the following structure. Initially, all attacks are launched one by one into each of x vulnerable targets until the threshold level

$$d(x) = \min_{i \geq 1} \left\{ i | \rho(x|\theta) (1-p)^i < 1 \right\} \tag{A9}$$

is reached in each of them or the attack resources are exhausted. Afterwards, the attacks are added one by one to non-vulnerable target until there is an attack on each of them. Then, attacks are added one by one until each of the vulnerable targets has $d(x) + 1$ attacks on each of them, then back to non-vulnerable targets until each has at least 2 attacks, etc. This "fill and switch" process stops when the attacker runs out of resources. After the process is complete, the attacker will have to uniformly randomize the distributions of attacks over sets of targets with the same sign: otherwise, the uniform distribution of protection by the defender would not be a best response. The threshold $\overline{a}(x)$ in the statement of Proposition 1 is a function of both x (via $d(x)$) and the total number of attacks available, a.

Lemma A4 *(i) Let $\pi(x) = (a_i, i = 1, 2, ..., n)$ be an optimal strategy . Then $|a_{i_1} - a_{i_2}| \leq 1$ when the signals at targets i_1, i_2 have the same sign.*

(ii) Let $\pi(x) = (a_i, i = 1, 2, ..., n)$ be a strategy, $0 < x < n$, u^- be the number of attacks in some vulnerable target, u^+ be the number of attacks in some protected target, and $d = d(x)$ is defined by formula (A9). Then, if $u^- - u^+ > d(x)$ or, if $u^+ \geq 1$ and $u^- - u^+ < d(x) - 1$, then strategy π is not optimal, or, equivalently, if π is optimal, and $u^+ = 0$, then $1 \leq u^- \leq d(x)$, and if $u^+ \geq 1$, then $u^- - u^+ = d(x)$ or $d(x) - 1$.

Proof. (i) Let J be a subset of targets, and recall that $C_j = 1$ when target j is destroyed. The conditional independence of testing and attacks' successes, and total probability formula imply the following formula for the conditional probability of the destruction of a particular target with $u \geq 1$ attacks

$$P(C|a, F) = P(C|a, T = 0) P(T = 0|F) = p(a) P(T = 0|F),$$

where F is any event generated by testing (signals).

Using the above formula and the definitions of $\rho(x)$, $p^-(x)$ and $p^+(x)$, we have:

$$\begin{aligned} P(C|a, S = 1, x) &= p^+(x) p(a), \\ P(C|a, S = 0, x) &= \rho(x) p^+(x) p(a). \end{aligned} \tag{A10}$$

Suppose that the statement is not true and let us say $a_{i_1} = a, a_{i_2} = j, a - j \geq 2$ and $S_{i_1} = S_{i_2} = 1$. The concavity of function $p(\cdot)$ implies that $p(a+1) + p(j-1) > p(a) + p(j)$. Then, using the formulas in (A10), we have

$$P(C=1|a+1, S=1, x) + P(C|j-1, S=1, x) > P(C|a, S=1, x) + P(C|j, S=1, x) \tag{A11}$$

Thus, $v(\pi|x)$ is not maximized and our initial strategy is not optimal. The proof for $S_{i_1} = S_{i_2} = 0$ is similar with $p^+(x)$ replaced by $p^-(x) = \rho(x)p^+(x)$.

(ii) Let $d(x) = d$. We will show that if $u^- - u^+ > d$ for some pair of vulnerable and protected targets, then a transfer of one attack from a vulnerable target to a protected target will increase the strategy's value. Similarly, if $u^+ \geq 1$ and $u^- - u^+ < d-1$ for such pair, then the inverse transfer will increase the value. As $\theta > \frac{1}{2}$, $\rho(x) > 1$ for $0 < x < n$, and hence $u^- \geq u^+$. Let $u^- = a, u^+ = j, P(\cdot|N=x) = P(\cdot|x)$, and denote the incremental utilities for vulnerable and protected targets as

$$\Delta C^-(u|x) = P(C|u+1, S', x) - P(C|u, S'^+(j|x)) = P(C|j+1, S, x) - P(C|j, S, x).$$

Then, formula (A10) implies that their difference for $0 \leq j \leq i$ is

$$\begin{aligned} \Delta(u,j|x) &= \Delta C^-(u|x) - \Delta C^+(j|x) \\ &= p\,(1-p)^j\, p^+(x)[\rho(x)\,(1-p)^{u-j} - 1]. \end{aligned}$$

The definition of $d = d(x)$ in (A9) implies that $\Delta(u,j|x)$ is positive if $j = 0, u < d$, or if $j \geq 1, u - j < d$. Similarly, $\Delta(u,j|x)$ is negative if $j = 0, u \geq d$, or if $j \geq 1, u - j \geq d$. These inequalities imply the claim. Also, note that if $p = 1$, then $d(x) = 1$ for all $0 < x < n$, and if p is decreasing to zero, then $d(x)$ tends to infinity. ■

This concludes the proof of Proposition 1. ■

Proof of Proposition 2

Take the critical ratio $\rho_{n,r}(x|\theta)$ defined by (A7) in the proof of Lemma A3. For each r, use (A7) and induction on n to show that $\rho_{n,r}(x|\theta)$ is an increasing function of θ for each x, $x \leq n-1$. Indeed, $\rho_{n,r}(x|\theta)$ is monotonically increasing in $\rho_{n-1,r}(x|\theta)$. Then, the induction step completes the argument.

Now, take any θ_1, θ_2 such that $\theta_1 < \theta_2$. As $\rho_{n,r}(x|\theta_1) < \rho_{n,r}(x|\theta_2)$, for the thresholds $d(x|\theta_1)$ and $d(x|\theta_2)$ defined by (A9), one has $d(x|\theta_1) \leq d(x|\theta_2)$. This means that attacks are more temporally concentrated with θ_2 than with θ_1, i.e., $\overline{a}(x|\theta_1) \leq \overline{a}(x|\theta_2)$. ■

Appendix C – Additional Institutional Details

In this section, we describe several additional details about the context. It is worth noting, as highlighted by Mansfield (2016), taxation by the Taliban is not perfectly uniform. In areas where the Taliban are not sufficiently powerful, the amount collected may be reduced (reflecting a local-level bargain with tribal leaders). Microlevel data on tax collection do not exist. Similarly we do not observe actual tax transfers after redistribution. Given the approach we detail later, we anticipate that each of these dynamics would bias our estimated effects downward.

It is also relevant to clarify that growth in opium cultivation in Afghanistan primarily benefits the Taliban. Unlike legal commodities in different conflicts—for example, oil in Colombia (Dube and Vargas, 2013; Wright, 2016) or bananas in the Philippines (Crost and Felter, 2020)—where both rebels and the central government can build capacity through taxation and/or extortion, illicit commodities primarily benefit non-state actors willing to usurp government regulation, including opium in Afghanistan, illegal timber sales in Burma and Cambodia, coca in Colombia (Estancona, 2021), and diamonds in Liberia and Sierra Leone (Le Billon, 2001). As such, we expect any capital shocks associated with increased suitability for opium production will enable the Taliban, specifically, to engage in more intense and sophisticated attacks.

Additionally, revenue tied to the opium trade is not the Taliban's only income source. The Taliban also benefited from external support—spaces to train Pakistan, weapons from Iran, and donations from Gulf states—as well as illegal mining activities. We focus on the opium trade for several reasons. First, opium was an essential source of funding during the period of study, with estimated revenue ranging from \$100-400 million USD annually. Although it is difficult to confirm the value of other revenue sources, the United Nations estimates that opium proceeds account for between 25-50% of the Taliban's annual income while the United States military assesses opium's income share at more than 60%. Second, opium production is more readily observable than other revenue sources and potential taxes can be matched more easily at the district-level, enabling us to study microlevel changes in resource endowments. As an agricultural commodity, we are also able to take advantage of how plausibly exogenous agronomic conditions impact local productivity.

Appendix D – Research Design

Appendix D1 – Additional measures

We gather a wealth of additional information about potentially relevant agricultural, demographic, and geographic factors. To evaluate development assistance, we leverage declassified data from the Commander's Emergency Response Program (CERP), a military-led scheme for small-scale development projects. We use these data to track overall aid provision as well as agricultural and irrigation projects specifically. We also gather data collected by the Food and Agriculture Organization (FAO) prior to the US invasion documenting irrigated sites.

We use these data to classify districts by irrigation intensity. We also follow the historical approach in Gehring, Langlotz and Kienberger (2019) to map Taliban territorial control using an archive produced in Dorronsoro (2005). We measure variability in terrain ruggedness using the raster data provided in Shaver, Carter and Wangyal Shawa (2019), following the approach in Carter, Shaver and Wright (2019). We also aggregate spatial information on languages spoken to calculate the percentage of settlements in a given district that are Pashtun. To do this, we rely on settlement data compiled by the Afghanistan Information Management Service, Central Statistics Office, United States Agency for International, and Yale University. We also follow the approach in Wigton-Jones (2021), using road segment completion data from the Afghanistan Information Management Service to measure market access.

Appendix D2 – Covariates in the main specification and robustness checks

In the manuscript, we briefly introduce the covariates in the main specification. We provide additional details here. To account for potential concerns about spatial correlation in production and local prices that varies over time, we include price zone-specific time trends. In addition, we allow for variation in these time trends by irrigation intensity, using data on the extent of irrigation infrastructure constructed prior to 2001, when the Taliban was removed from power. It is also important to note that the base terms for these quantities—a district's location in one of the six regional price zones or classification of available pre-invasion infrastructure—are captured in our district fixed effects.

To supplement these measures, we also include information about small-scale development programs led by security forces during the growing season and time-varying measure of market access. We incorporate these measures of development assistance because the location and scale of these projects may be correlated with opium productivity. That is, conditional on observing fluctuations in agronomic suitability, government forces strategically reallocate aid projects. If these aid projects effectively reduce opium production (via 'alternative livelihoods' programming) and enhance local employment, our main estimates may be biased towards zero since the introduction of these programs may raise reservation wages, attenuating any mobilization effects of potential revenue. It is also possible, however, that rebels are able to capture some of the rents associated with aid projects. This could bias our estimates upward, since changes in suitability are positively associated with a second, otherwise unmeasured source of revenue. We also take seriously the potential concern that changes in market access induced by expansion of the Ring Road highway network—championed as a centerpiece of Afghan reconstruction and development—may have significantly reduced transportation costs associated with moving refined opium products to regional markets or across international borders. This, as Wigton-Jones (2021) finds, may have significantly increased the amount of land under cultivation while also reducing the costs of relocating for rebel forces after combat engagements.

Our robustness checks incorporate additional covariates in the main specification. We begin by introducing a measure of terrain ruggedness following the approach in Shaver, Carter and Wangyal Shawa (2019), exploiting granular raster data of Afghanistan's topography. Since the baseline effect of terrain is already absorbed in our unit fixed effects, we introduce a ruggedness time trend. Terrain features may matter since geographic variability can create ideal locations to produce illicit crops while hiding from state forces. These areas may also create ideal locations from which to coordinate and launch attacks.

The human terrain of a district might also ease constraints on the production of violence. We attempt to address this dynamic by incorporating a time trend for co-ethnic density in each district. Although a comprehensive map of population distribution of ethnic groups in Afghanistan is not available, we use georeferenced settlement-level data about languages spoken to identify Pashto-speaking settlements, which we then use to measure the percentage of a district's population that speaks Pashto. From this, we can make inferences about the location and density of Pashtuns, the ethnic group that during this period has split allegiances between the government, with a Pashtun holding the presidency, and the Taliban, which is composed predominantly by Pashtuns. Co-ethnic ties may ease the ability, on average, for Taliban forces to promote cultivation, collect post-harvest taxes, and engage in intelligence gathering. Relatedly, we also incorporate a time trend tied to a historical measure of the Taliban's consolidation of control at the end of 1996, when they initially seized control of Kabul and, with it, the central government.

We also supplement the main specification with several measures of coercive shows of force and intimidation by insurgents during the planting and growing seasons. Insurgents may use these violent and non-violent tactics to compel civilians to engage in opium cultivation, such as killings of government collaborators and the posting of 'night letters' and other non-lethal shows of force. We also incorporate measures of counterinsurgent operations that may be closely tied to operational planning (by the Taliban), details of the internal organization of local forces, and the ability to gather information about combat vulnerabilities. These measures include events where security forces engaged in search and seize operations, where they gain access to actionable intelligence about Taliban activities and resources as well as detention of suspected fighters and collaborators.

Appendix D3 – Details on temporal clustering measure

In Section 5.2, we describe evidence linking opium revenue and clustering of attacks in time and, separately, space. For both methods, we pool information across the primary attack types, allowing us to measure combat clustering generally. We begin by detailing the method for measuring temporal clustering. The central intuition of this approach is to use randomization inference to better understand the degree of temporal clustering we observe in insurgent combat operations. We identify the hour of each attack within a given district-year (fighting season). We then reshuffle the hour vector and compare the empirical distribution to the randomly reshuffled vector. This process is repeated many times per district-year. The result of the technique is a single likelihood parameter, which we call a p-value, for each unit of observation. We estimate these parameters for district-years with a minimum of five conflict events.[2] Higher p-values indicate that the distribution of rebel attacks by hour cannot be distinguished from randomness. Lower p-values reveal attack patterns that are more easily differentiated from randomness; i.e., they are more clustered. By design, this approach cannot tell us whether a given event is randomly timed. Rather, we study whether fighting during a period of time (e.g., a fighting season) is more or less random. The approach is executed in several steps.

1. Fit a local polynomial regression to the observed distribution of violence by hour. We specify a conservative bandwidth of 1. This empirical distribution of fitted values is stored.[3]
2. Identify the sequence of district-hours during which combat engagements occur. For each district-hour, we know the sum of the number of attacks.
3. Randomly shuffle the sequence above. This is equivalent to a randomization or permutation test.
4. Fit a local polynomial regression to the randomly shuffled distribution of violence by hour. The simulated distribution of fitted values is stored.
5. Execute the bootstrap Kolmogorov-Smirnov test composed of four elements.
    (a) Compute the T^{KS}_{dfi} for the fitted values of the empirical and simulated distributions, where
    $$T^{KS}_{dfi} = \left(\frac{n_1 n_0}{n}\right)^{\frac{1}{2}} \sup_{y \in \mathbb{R}} \left|F_{1,n_1}(y) - F_{0,n_1}(y)\right|.$$
    (b) Resample observations with replacement from observed and simulated distributions. Split the resampled set into two distributions; calculate and store $T^{KS}_{dfi,b}$.
    (c) Repeat prior two steps 1,000 times.
    (d) Calculate and store the likelihood parameter of the tests as $\sum_{b=1}^{1000} \frac{1\{T^{KS}_{dfi,b} > T^{KS}_{dfi}\}}{1{,}000}$, where the numerator is an indicator function.
6. Repeat steps 2 through 5 1000 times. Evaluate the central tendency (mean) of the likelihood parameters.
7. Replace zero values with the minimum observed non-zero rank value and calculate the log.

---

[2] We set the lower threshold at five events to ensure convergence of the simulations. A conflict vector that is too short (i.e., fewer than five) does not permit sufficient randomization when the hour vector is reshuffled. Our results are highly consistent if we raise this threshold upward.

[3] Some conflict events lack a time stamp ($\sim$3%). Because we cannot assign these events an hour, they are excluded from the calculation of the empirical distribution.

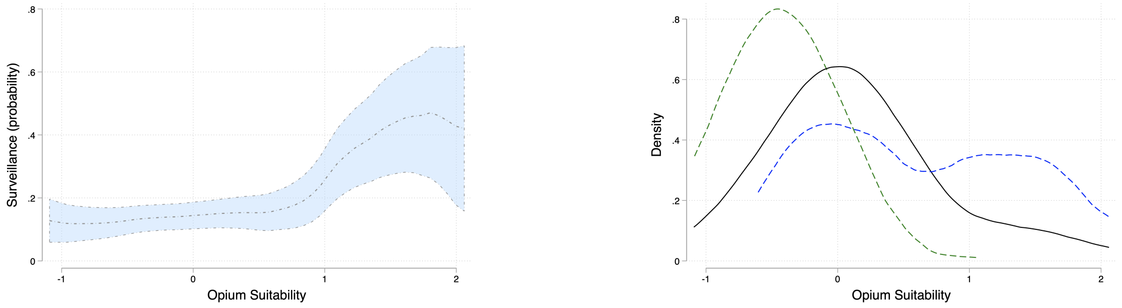
Appendix D4 – Details on spatial clustering measure

In this subsection we detail a test of spatial clustering in combat patterns. This section supplements Section 5.2, where we introduce evidence linking opium revenue and clustering of attacks.

The central inferential challenge for studying spatial dynamics is identifying the set of counterfactual locations that could have been attacked but were not. If, for example, the set of locations that could be targeted was a known quantity and did not change over time, it would be possible to identify clustering among the finite set of targets. In Afghanistan, this quantity is unknown and even if it was recoverable from the highly classified Blue Force Tracker database, the number of potential targets and their location is endogenous to conflict dynamics. To identify relevant elements of the grid, we restrict our analysis only to grid cells that see some recorded activity during the entire conflict (extensive margin). This grid selection rule indicates that at some point during the war, there was a relevant target present within this cell (which may or may not have been a target of prior or later operations). To reduce potential concerns about endogenous changes in the location of targets due to spatial attack patterns, we keep this grid sample fixed over time. We then calculate a dispersion index, following Perry and Mead (1979).

Appendix E – Supplemental Figures and Results

Figure A-1: Surveillance Operations by Rebels Associated With Exogenous Opium Suitability Measure

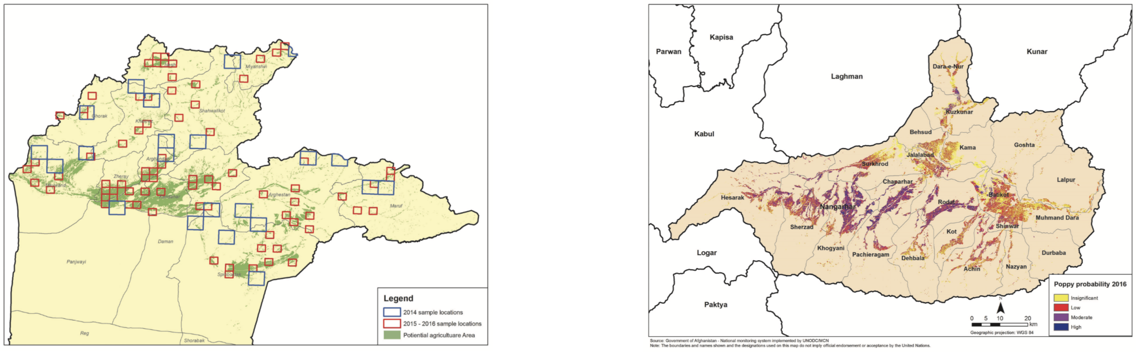


(a) Opium Suitability and Surveillance

(b) Opium Suitability by Price Zone

Notes: Panel (a) depicts a local polynomial regression linking exogenous variation in opium suitability and the location of rebel-led surveillance operations at the beginning of the study sample. We use this cross-sectional measure in our heterogeneous effects. See Equation 4. Panel (b) depicts suitability in 2006 across price zones. The green line indicates a very weak distribution of suitability in the central price region anchored by Kabul. The blue line indicates a robust suitability distribution for the southern price region, anchored by Hilmand and Kandahar.

Figure A-2: UNODC Methodology for Estimating Annual District Drug Production



(a) Sampling Satellite Imagery

(b) Impute Production from Imagery/Field Obs.

Notes: Methodological figures and details drawn from the 2016 UNODC-Afghanistan Drug Report. Panel (a) demonstrates the sampling design used when acquiring high resolution satellite imagery (location: Kandahar). Panel (b) illustrates the subsequent production estimation, which combines low and high resolution imagery (location: Nangahar).

Table A-2: Robust Association Between Exogenous Suitability Parameter and Insurgent Intelligence Gathering

	(1) Baseline	(2) Suitability by Price Zone	(3) + Full Covariates	(4) + Province Fixed Effects
Opium Suitability	0.089**			
	(0.032)			
Suit. (Kabul Price Zone)		-0.206*	-0.252**	-0.282**
		(0.083)	(0.085)	(0.080)
Suit. X Nangarhar Price Zone		0.265**	0.192†	0.328*
		(0.098)	(0.115)	(0.129)
Suit. X Kunduz Price Zone		0.031	0.074	0.131
		(0.131)	(0.128)	(0.157)
Suit. X Balkh Price Zone		0.276**	0.292**	0.352**
		(0.095)	(0.098)	(0.095)
Suit. X Hilmand/Kandahar Price Zone		0.322**	0.313**	0.445**
		(0.106)	(0.102)	(0.123)
Suit. X Hirat Price Zone		0.264*	0.253*	0.328**
		(0.104)	(0.103)	(0.101)
SUMMARY STATISTICS				
Outcome Mean	0.176	0.176	0.176	0.176
Outcome SD	0.381	0.381	0.381	0.381
MODEL PARAMETERS				
Price Zone Specific Effects	No	Yes	Yes	Yes
Irrigation Intensity	No	No	Yes	Yes
Dev aid: Ag/Irrigation	No	No	Yes	Yes
Market Access	No	No	Yes	Yes
Terrain Rug.	No	No	Yes	Yes
Coethnicity	No	No	Yes	Yes
Taliban Control	No	No	Yes	Yes
Coercion	No	No	Yes	Yes
COIN Ops	No	No	Yes	Yes
Province Fixed Effects	No	No	No	Yes
MODEL STATISTICS				
No. of Observations	398	398	398	398
No. of Clusters	398	398	398	398
R^2	0.031	0.193	0.272	0.362

Notes: Outcome of interest is intelligence gathering by rebels at the start of the sample (2006, =1). Column headings indicate model specification. Additional details are shown in the table rows below estimation results. Column 1 is a simple bivariate regression, suggesting that exogenous opium suitability is strongly correlated with surveillance in the cross section. Column 2 allows the estimated effects of suitability to vary by region. Consistent with the suitability distributions depicted in Figure A-1 (b), the Kabul price zone registered very weak suitability in 2006 despite having a fairly significant number of districts with surveillance activities. This yields a negative base term. The positive marginal effects indicate significant differences across four of the other price zones. These differences are robust to a fully saturated specification following Table 1 (column 3) as well as a demanding province fixed effects approach (column 4). All regressions include district and year fixed effects as well as controls as specified under model parameters. Heteroskedasticity robust standard errors are reported in parentheses. Symbols indicate † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

In Table A-9, we present an alternative estimation strategy that leverages a balanced panel approach to the results presented in Table 3. Because the clustering parameters are only defined conditional on violence reaching a specific threshold within the fighting season (i.e., at least five events), the baseline approach uses an unbalanced panel. We adjust this approach by replacing missing values with 0 and adding an indicator variable (for replacement) to the regression. This indicator effectively turns on for any district-year observation when the observation would otherwise be missing from the panel. This allows our covariate-specific time trends to have the same interpretation as the other balanced panel outcomes. Notice that the results are largely unaffected by the correction, with clustering

Table A-3: Impact of Rebel Capacity and Income Volatility on Combat Activity

	(1)	(2)	(3)	(4)
Opium Suitability	$0.560^{\dagger}$	$0.595^{\dagger}$	$0.493^{\dagger}$	$0.448^{\dagger}$
	(0.334)	(0.356)	(0.294)	(0.241)
Suit. × High Income Volatility (med.)	-0.288			
	(0.323)			
Suit. × High Income Volatility (mean)		-0.340		
		(0.341)		
Suit. × High Income Volatility (75 ptile)			-0.155	
			(0.305)	
Suit. × High Income Volatility (90 ptile)				0.025
				(0.439)
SUMMARY STATISTICS				
Outcome Mean	0.352	0.352	0.352	0.352
Outcome SD	1.227	1.227	1.227	1.227
MODEL PARAMETERS				
District Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Price Zone Trends	Yes	Yes	Yes	Yes
Irrigation Intensity Trends	Yes	Yes	Yes	Yes
Dev aid: Ag/Irrigation	Yes	Yes	Yes	Yes
Market Access	Yes	Yes	Yes	Yes
MODEL STATISTICS				
No. of Observations	3582	3582	3582	3582
No. of Clusters	398	398	398	398
R^2	0.581	0.582	0.580	0.580

Notes: Outcome of interest is overall combat activity (equivalent to Column 1 of Table 1. The quantity of interest is opium suitability interacted with various thresholds of income volatility calculated using potential revenue among producers. All regressions include district and year fixed effects as well as controls as specified under model parameters. Heteroskedasticity robust standard errors clustered by district are reported in parentheses. Symbols indicate † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

continuing to increase in revenue, especially with respect to areas with historical intelligence gathering by rebels.

Table A-4: Robust Association Between Various Endogenous Opium Revenue Measures and Exogenous Suitability Parameter

	(1) Revenue	(2) Revenue, yield adjusted	(3) Revenue, regional price/yield adj
Opium Suitability	2.408**	4.099**	4.125**
	(0.598)	(0.976)	(0.961)
SUMMARY STATISTICS			
Outcome Mean	8.256	13.567	13.194
Outcome SD	14.041	21.803	21.208
MODEL PARAMETERS			
District Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Price Zone Trends	Yes	Yes	Yes
Irrigation Intensity Trends	Yes	Yes	Yes
Dev aid: Ag/Irrigation	Yes	Yes	Yes
Market Access	Yes	Yes	Yes
MODEL STATISTICS			
No. of Observations	3582	3582	3582
No. of Clusters	398	398	398
R^2	0.794	0.778	0.776

Notes: Outcome of interest varies by column and is indicated in the column heading. The quantity of interest is opium suitability. Column 1 depicts log output in hectares by log of the simple average national price. Column 2 adjusts the outcome in column 1 using an annual yield calibration weight, allowing us to convert hectares under production to estimation kilograms. Column 3 allows for regional yield adjustment as well as price zone-by-year price changes. All regressions include district and year fixed effects as well as controls as specified under model parameters. Heteroskedasticity robust standard errors clustered by district are reported in parentheses. Symbols indicate † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

Table A-5: Impact of Rebel Capacity and Surveillance on Combat Outcomes With Additional Covariates

	(1) Combat	(2) Direct Fire	(3) IED Explosion	(4) Indirect Fire
Panel A: Baseline Effects with additional covariates				
Opium Suitability	$0.348^{\dagger}$	$0.272^{\dagger}$	0.074^{**}	0.001
	(0.177)	(0.157)	(0.023)	(0.006)
SUMMARY STATISTICS				
Outcome Mean	0.352	0.223	0.071	0.058
Outcome SD	1.227	0.963	0.224	0.228
MODEL STATISTICS				
No. of Observations	3582	3582	3582	3582
No. of Clusters	398	398	398	398
R^2	0.621	0.567	0.664	0.598
Panel B: Heterogeneous Effects with additional covariates				
Opium Suitability	0.038	0.001	0.029^{*}	0.007
	(0.049)	(0.039)	(0.012)	(0.007)
Suit. × Rebel Surveillance	$0.634^{\dagger}$	$0.554^{\dagger}$	0.092^{*}	-0.012
	(0.352)	(0.311)	(0.045)	(0.011)
SUMMARY STATISTICS				
Outcome Mean	0.352	0.223	0.071	0.058
Outcome SD	1.227	0.963	0.224	0.228
ADDITIONAL MODEL PARAMETERS				
Terrain Rug. Trends	Yes	Yes	Yes	Yes
Coethnicity Trends	Yes	Yes	Yes	Yes
Taliban Control Trends	Yes	Yes	Yes	Yes
Coercion	Yes	Yes	Yes	Yes
COIN Ops	Yes	Yes	Yes	Yes
MODEL STATISTICS				
No. of Observations	3582	3582	3582	3582
No. of Clusters	398	398	398	398
R^2	0.629	0.576	0.669	0.598

Notes: Outcome of interest varies by column and is indicated in the column heading. The quantity of interest is opium suitability. All regressions include district and year fixed effects as well as controls as specified under model parameters. Heteroskedasticity robust standard errors clustered by district are reported in parentheses. Symbols indicate † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

Table A-6: Impact of Rebel Capacity and Surveillance on Combat Innovation, Coordination, and Complexity With Additional Covariates

	(1) Tactical Innovation	(2) Deceptive Tech	(3) Unit Breach	(4) Complex Target
Panel A: Baseline Effects with additional covariates				
Opium Suitability	0.004*	0.005**	0.001*	0.033*
	(0.002)	(0.002)	(0.000)	(0.013)
SUMMARY STATISTICS				
Outcome Mean	0.003	0.004	0.001	0.014
Outcome SD	0.014	0.021	0.004	0.085
MODEL STATISTICS				
No. of Observations	3582	3582	3582	3582
No. of Clusters	398	398	398	398
R^2	0.291	0.388	0.240	0.446
Panel B: Heterogeneous Effects with additional covariates				
Opium Suitability	-0.001	0.002	0.000	0.010
	(0.001)	(0.002)	(0.000)	(0.006)
Suit. × Rebel Surveillance	0.010*	0.007*	0.001	0.047*
	(0.004)	(0.003)	(0.001)	(0.022)
SUMMARY STATISTICS				
Outcome Mean	0.003	0.004	0.001	0.014
Outcome SD	0.014	0.021	0.004	0.085
ADDITIONAL MODEL PARAMETERS				
Terrain Rug. Trends	Yes	Yes	Yes	Yes
Coethnicity Trends	Yes	Yes	Yes	Yes
Taliban Control Trends	Yes	Yes	Yes	Yes
Coercion	Yes	Yes	Yes	Yes
COIN Ops	Yes	Yes	Yes	Yes
MODEL STATISTICS				
No. of Observations	3582	3582	3582	3582
No. of Clusters	398	398	398	398
R^2	0.304	0.391	0.241	0.455

Notes: Outcome of interest varies by column and is indicated in the column heading. The quantity of interest is opium suitability. All regressions include district and year fixed effects as well as controls as specified under model parameters. Heteroskedasticity robust standard errors clustered by district are reported in parentheses. Symbols indicate † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

Table A-7: Impact of Rebel Capacity and Surveillance on Attack Clustering With Additional Covariates

	(1) Temporal	(2) Temporal (TWFE)	(3) Spatial	(4) Spatial (TWFE)
Panel A: Baseline Effects with additional covariates				
Opium Suitability	-0.463	-0.290	429.771*	313.064†
	(0.402)	(0.484)	(189.317)	(169.174)
Summary Statistics				
Outcome Mean	-7.048	-7.110	343.140	350.022
Outcome SD	4.355	4.368	1072.541	1083.435
Model Statistics				
No. of Observations	1467	1435	1467	1435
No. of Clusters	266	234	266	234
R^2	0.129	0.391	0.241	0.623
Panel B: Heterogeneous Effects with additional covariates				
Opium Suitability	0.316	0.515	9.481	31.342
	(0.290)	(0.360)	(73.625)	(70.779)
Suit. × Rebel Surveillance	-1.537*	-1.511†	825.972*	528.713†
	(0.697)	(0.813)	(324.792)	(274.256)
Summary Statistics				
Outcome Mean	-7.048	-7.110	343.140	350.022
Outcome SD	4.355	4.368	1072.541	1083.435
Additional Model Parameters				
Terrain Rug. Trends	Yes	Yes	Yes	Yes
Coethnicity Trends	Yes	Yes	Yes	Yes
Taliban Control Trends	Yes	Yes	Yes	Yes
Coercion	Yes	Yes	Yes	Yes
COIN Ops	Yes	Yes	Yes	Yes
Model Statistics				
No. of Observations	1467	1435	1467	1435
No. of Clusters	266	234	266	234
R^2	0.155	0.396	0.305	0.631

Notes: Outcome of interest varies by column and is indicated in the column heading. The quantity of interest is opium suitability. All regressions include year fixed effects as well as controls as specified under model parameters. Even columns include district fixed effects. Heteroskedasticity robust standard errors clustered by district are reported in parentheses. Symbols indicate † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

Table A-8: Impact of Rebel Capacity and Surveillance on Combat Losses and Casualties With Additional Covariates

	(1) Disrupt: Govt	(2) Casualties: Govt	(3) Disrupt: Coal	(4) Casualties: Coal
Panel A: Baseline Effects with additional covariates				
Opium Suitability	0.045*	0.018	0.057**	0.030*
	(0.018)	(0.012)	(0.021)	(0.013)
SUMMARY STATISTICS				
Outcome Mean	0.079	0.060	0.031	0.017
Outcome SD	0.255	0.202	0.142	0.083
MODEL STATISTICS				
No. of Observations	3582	3582	3582	3582
No. of Clusters	398	398	398	398
R^2	0.650	0.632	0.486	0.483
Panel B: Heterogeneous Effects with additional covariates				
Opium Suitability	0.006	-0.010	0.019*	0.005
	(0.009)	(0.008)	(0.010)	(0.005)
Suit. × Rebel Surveillance	0.081*	0.058**	0.078†	0.051*
	(0.034)	(0.021)	(0.042)	(0.025)
SUMMARY STATISTICS				
Outcome Mean	0.079	0.060	0.031	0.017
Outcome SD	0.255	0.202	0.142	0.083
ADDITIONAL MODEL PARAMETERS				
Terrain Rug. Trends	Yes	Yes	Yes	Yes
Coethnicity Trends	Yes	Yes	Yes	Yes
Taliban Control Trends	Yes	Yes	Yes	Yes
Coercion	Yes	Yes	Yes	Yes
COIN Ops	Yes	Yes	Yes	Yes
MODEL STATISTICS				
No. of Observations	3582	3582	3582	3582
No. of Clusters	398	398	398	398
R^2	0.653	0.635	0.494	0.494

Notes: Outcome of interest varies by column and is indicated in the column heading. The quantity of interest is opium suitability. All regressions include district and year fixed effects as well as controls as specified under model parameters. Heteroskedasticity robust standard errors clustered by district are reported in parentheses. Symbols indicate † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

Table A-9: Impact of Rebel Capacity and Surveillance on Attack Clustering

	(1)	(2)	(3)	(4)
	Temporal (Unbalanced)	Temporal (Balanced)	Spatial (Unbalanced)	Spatial (Balanced)
Panel A: Baseline Effects				
Opium Suitability	-0.349	-0.246	346.963†	210.113*
	(0.510)	(0.266)	(198.850)	(103.826)
SUMMARY STATISTICS				
Outcome Mean	-7.110	-2.887	350.022	147.452
Outcome SD	4.368	4.447	1083.435	705.583
MODEL STATISTICS				
No. of Observations	1435	3582	1435	3582
No. of Clusters	234	398	234	398
R^2	0.387	0.729	0.599	0.546
Panel B: Heterogeneous Effects				
Opium Suitability	0.471	0.170	37.973	28.158
	(0.361)	(0.151)	(67.666)	(26.824)
Suit. × Rebel Surveillance	-1.529†	-1.257†	576.607†	549.840*
	(0.834)	(0.689)	(315.301)	(268.170)
SUMMARY STATISTICS				
Outcome Mean	-7.110	-2.887	350.022	147.452
Outcome SD	4.368	4.447	1083.435	705.583
MODEL PARAMETERS				
District Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Price Zone Trends	Yes	Yes	Yes	Yes
Irrigation Intensity Trends	Yes	Yes	Yes	Yes
Dev aid: Ag/Irrigation	Yes	Yes	Yes	Yes
Market Access	Yes	Yes	Yes	Yes
MODEL STATISTICS				
No. of Observations	1435	3582	1435	3582
No. of Clusters	234	398	234	398
R^2	0.391	0.731	0.609	0.560

Notes: Outcome of interest varies by column and is indicated in the column heading. The quantity of interest is opium suitability. All regressions include district and year fixed effects as well as controls as specified under model parameters. Even columns exploit alternative balanced panel estimation approach. Heteroskedasticity robust standard errors clustered by district are reported in parentheses. Symbols indicate † $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

References

Blackett, D. W. 1958. "Pure Strategy Solutions of Blotto Games." *Naval Research Logistics Quarterly* 5(2):107–109.

Borel, Emil. 1921. "La theorie du jeu les equations integrales a noyau symetrique. Comptes Rendus del Academie, 173, 1304-1308; English translation by Savage, L. (1953). The theory of play and integral equations with skew symmetric kernels." *Econometrica* 21:97–100.

Carter, David B., Andrew C. Shaver and Austin L. Wright. 2019. "Places to Hide: Terrain, Ethnicity, and Civil Conflict." *The Journal of Politics* 81(4):1446–1465.

Crost, Benjamin and Joseph H Felter. 2020. "Export crops and civil conflict." *Journal of the European Economic Association* 18(3):1484–1520.

Dorronsoro, Gilles. 2005. *Revolution unending: Afghanistan, 1979 to the present.* Columbia University Press.

Dube, Oeindrilla and Juan Vargas. 2013. "Commodity Price Shocks and Civil Conflict: Evidence from Colombia." *Review of Economic Studies* 80(4):1384–1421.

Estancona, Chelsea L. 2021. "Rebel Primary Commodity Markets, Price Shocks, and Supplier Victimization." *International Studies Quarterly* 65(4):1111–1123.

Gehring, Kai, Sarah Langlotz and Stefan Kienberger. 2019. "Stimulant or Depressant? Resource-Related Income Shocks and Conflict.".

Golman, Russell and Scott E. Page. 2009. "General Blotto: games of allocative strategic mismatch." *Public Choice* 138(3):279–299.

Konrad, Kai A. and Dan Kovenock. 2009. "Multi-battle contests." *Games and Economic Behavior* 66(1):256–274.

Kovenock, Dan and Brian Roberson. 2012. "Conflicts with Multiple Battlefields," in Michelle R. Garfinkel and Stergios Skaperdas (eds.). Oxford Handbook of the Economics of Peace and Conflict, Oxford University Press, New York. pp. 503-531.".

Kress, Moshe. 2012. "Modeling Armed Conflicts." *Science* 336(6083):865–869.

Le Billon, Philippe. 2001. "The Political Ecology of War: Natural Resources and Armed Conflict." *Political Geography* 20(5):561–584.

Mansfield, David. 2016. *A State Built on Sand: How Opium Undermined Afghanistan.* Oxford: Oxford University Press.

Perry, J. N. and R. Mead. 1979. "On the Power of the Index of Dispersion Test to Detect Spatial Pattern." *Biometrics* 35(3):613–622.

Powell, Robert. 2007. "Defending against Terrorist Attacks with Limited Resources." *American Political Science Review* 101(3):527–541.

Shaver, Andrew, David B. Carter and Tsering Wangyal Shawa. 2019. "Terrain ruggedness and land cover: Improved data for most research designs." *Conflict Management and Peace Science* 36(2):191–218.

Wigton-Jones, Evan. 2021. "The unintended harms of infrastructure: Opium and road construction in Afghanistan." *Journal of Comparative Economics* 49(2):405–424.

Wright, Austin L. 2016. "Economic Shocks and Rebel Tactics: Evidence from Colombia." *HiCN paper #232* .