Rebel Capacity, Intelligence Gathering, and the Timing of Combat Operations*

Konstantin Sonin, Jarnickae Wilson, and Austin L. Wright†

September 6, 2018

Abstract

Classic theories of counterinsurgency claim rebel forces execute attacks in an unpredictable manner to limit the government's ability to anticipate and defend against them. We study a model of combat and information-gathering during an irregular insurgency. We test empirical implications of the model using newly declassified military records from Afghanistan that include highly detailed information about rebel attacks and counterinsurgent operations, including close air support missions, bomb neutralizations, and covert government-led surveillance activity. Our conflict micro-data also include previously unreleased information about insurgent-led spy networks, where rebels monitor troop movement and military base activity, as well as military base infiltration and insider attacks. We couple these data with granular information on opium production and farmgate prices. Consistent with our simple theoretical model, we find that the capacity (wealth) of local rebel units influences the timing of their attacks. As rebels gather more resources, their attacks become temporally concentrated in a manner that is distinguishable from randomized combat. This main effect is significantly enhanced in areas where rebels have the capacity to spy on and infiltrate military installations. Taken together, these findings suggest economic shocks that increase the capacity of insurgents may influence the timing of rebel attacks through the acquisition of precise information about military weaknesses.

*For advice and comments, we thank Peter Deffebach, Francois Libois, Kyle Pizzey, Mehdi Shadmehr, Oliver Vanden Eynde, and Liam Wren-Lewis. This material is based in part upon work supported by the Pearson Institute for the Study and Resolution of Global Conflicts. Eli Berman, Ethan Bueno de Mesquita, and Jacob Shapiro are owed a particular debt of gratitude for their support of this and related projects. Maria Ballesteros, Connie Guo, Lana Kugli, Michelle Liang, Yonatan Litwin, Paul Soltys, and Julian Waugh provided excellent research assistance.

†Harris School of Public Policy, The University of Chicago, Chicago, IL 60637. Emails: ksonin@uchicago.edu, jarnicaew@uchicago.edu, and austinlw@uchicago.edu (corresponding author).

1 Introduction

In the 21st century, intrastate conflicts have replaced interstate ones as the main source of human loss and population displacement. Resource endowments shape how rebels recruit, retain, and deploy their fighters (Weinstein, 2007). As insurgents accumulate capacity, they can expand their control of strategic territory (Kalyvas, 2006). Territorial control influences how rebels treat civilians (Wood, 2014), whether civilians cooperate with rebels or collude with government forces (Condra and Shapiro, 2012), and impacts the ability of the government to engage in development and reconstruction (Sexton, 2016). Rebel capacity defines even the finest dimensions of internal warfare.

One central question that remains largely unexplored is whether rebel capacity influences *how* rebels fight. In particular, classic theories of insurgency note that the rebel's main advantage in an asymmetric conflict is the 'element of surprise' (Galula, 1964). If guerrillas aim to undermine their more powerful rivals, their attacks must be unpredictable and, as such, difficult for government forces to anticipate and thwart (Thompson, 1966; Powell, 2007a,b). Yet the strategic value of random combat may diminish as rebels accumulate resources. As their capacity grows, rebels may shift from guerrilla tactics to conventional warfare, where frontal assaults are less random and more costly to the rebel forces, but enable insurgents to consolidate territorial and economic control. Additional resources might also allow rebels to gather precise information about defensive weaknesses, including when troops and military installations are vulnerable to attack.

These dynamics motivate our theoretical argument. In our model of combat, rebels allocate scarce resources to conducting attacks at various times. The government also invests in defensive tactics to 'harden' (deter) against attacks at particular times when troops and bases might otherwise be vulnerable. Once the government allocates protection, the rebels obtain some noisy information about the protection (i.e., *when* targets are vulnerable). As government resources are scarce, information about the presence of protection at one point in time is simultaneously information about (the absence of) protection during other periods of the day. With higher quality information, rebels concentrate their attacks on times when targets are unprotected. This suggests a simple theoretical mechanism linking rebel capacity and the temporal randomness of insurgent attacks. With few resources, rebels can acquire relatively worse information (i.e., lower quality), and attacks appear to be more randomly timed (i.e., the timing of attacks exhibit higher entropy when the quality of information is lower).

We test the empirical implications of our model using newly declassified military records provided to us by the United States government. These records document hundreds of thousands of combat engagements in Afghanistan during Operation Enduring Freedom, including 28,678 instances of indirect fire attacks (typically rocket or mortar fire). These data are particularly valuable because they include within-day timestamps (often to the minute), which we collapse to the hour. This rich feature of the data allows us to study temporal clustering in attack patterns, consistent with a deviation from random attack timing. We quantify the randomness of combat timing by developing and implementing a novel extension of bootstrapping methods in statistics applied to the canonical Kolmogorov-Smirnov test. This method yields a likelihood parameter capturing the probability an empirical distribution of combat timing can be distinguished from a random attack pattern. We then study the association between temporal clustering and rebel capacity using granular data on opium revenue, which the insurgents consistently tax at a fixed rate (Peters, 2009).

We find evidence that the randomness of attacks significantly decreases as rebels accumulate more resources from the opium trade. As rebels accumulate fighting capacity, their attacks become temporally clustered around particular time windows in the day. This core result survives a number of robustness checks, including accounting for trends in rebel violence during the fighting, harvest, and planting seasons, which may influence the intensity of opium cultivation from year to year. The richness of our microdata also enables us to rule out additional concerns about the positive covariance between rebel capacity and strategic reallocation of government forces. We do this by accounting for variation in the location and timing of 5,410 close air support missions, 43,679 bomb neutralizations, 1,279 raids on insurgent safe houses, the capture and detention of 13,839 rebel fighters, and 776 covert government-led surveillance operations. Our military records also allow us to address potential concerns about the role of non-lethal coercion of opium production.

The following simple example (**Example 1**) illustrates the basic logic of our model, which captures a common feature of irregular warfare. Counterinsurgents deploy technology to protect bases from attack. Troop and vehicle movement as well as day-to-day operations make military installations more or less vulnerable during certain time slots (hours of the day). This was the case in Uruzgan province following the Afghan surge in 2010, when insurgents were able to monitor and assess weaknesses at one of the Coalition's forward operating bases.[1] Rebels had amassed an arsenal of rockets, which they deployed strategically at particular times during the day. Despite quick reaction times and a technological advantage, Coalition forces were unable to neutralize the insurgent threat and carefully timed attacks continued during the subsequent fighting season.

Example 1 Suppose that a rebel group considers launching attacks at two time intervals during the day, and that the government has resources to defend against attacks during only one period. The choice for the rebel group is whether to launch attacks at both times or attack during one period with a double strength. If a target is protected by the government when the attack occurs, the attack does not succeed. If the target is unprotected during the strike, the attack's chance of success is p, $0 < p < 1$; the probabilities are i.i.d..

Suppose that the government moves first: the equilibrium strategy is to randomize between the two time periods with probability 1/2. In the absence of any information, attacks

[1] See `https://tinyurl.com/y8e5x466`.

on different time slots result in one rebels' success with probability p; two attacks during the same time slot result in two successes with probability $\frac{1}{2}p^2$, and one success with probability $p(1-p)$. Now, suppose that after the defense is allocated, the rebels gathered information and the posterior probability that the first time window is vulnerable is $q > \frac{1}{2}$. Launching attacks at two different times results in the expected harm of p and the probability that at least one attack is successful of p. As $q > \frac{1}{2}$, if two attacks are launched, this is done in the first time window. The expected harm is $2qp^2 + 2qp(1-p)$, which exceeds p for any $q > \frac{1}{2}$, and the probability that at least one attack is successful is $qp^2 + 2qp(1-p)$, which exceeds p as long as $q > \frac{1}{2-p}$. Intuitively, rebels concentrate their rocket fire (or attacks more generally) when the attacking technology is inefficient (p is low) and intelligence about when the is target vulnerable is high quality (q is high).

The attacks become more temporally concentrated when information about target vulnerability is more precise. Now, suppose that rebels receive conditionally independent binary signals $s_i \in \{V, D\}$ such that $\Pr(s_i = V|V) = \Pr(s_i = D|D) = \theta \geq \frac{1}{2}$, larger θ corresponding to a more precise information. (These signals are not independent as if the first time slot is vulnerable, the second is not, and vice-versa.) Furthermore, suppose that $s_1 = V, s_2 = D$. Then $q = \frac{\theta^2}{\theta^2+(1-\theta)^2}$, which is increasing in θ, and a higher θ makes concentration of attacks more likely. Thus, when rebels have fewer resources and spend less on information gathering, their attacks become less concentrated. In Section 2, we analyze the equilibrium of this model for any number of time slots N, periods that are protected by the government disposal R, and rebel capacity (the number of possible attacks) A.[2]

Afghanistan provides an ideal setting to test the potential mechanisms that explain the

[2]The general setup might be familiar to a reader: a particular case of our model can been used to demonstrate the popular "Monty Hall paradox" (e.g., Gill, 2010). In the model with $N = 3$, $R = 2$, $p = 1$, $A = 1$, if the rebels' choice is known to the government and it does, without reallocating the defense, reveal that one of the non-chosen periods is protected, the optimal *ex ante* strategy for rebels is to switch away from their initially chosen time of attack. In the setup with partial information, our general case, the optimality of switching would depend on the parameters of the model.

relationship between rent extraction and the timing of insurgent attacks. To do this, we first develop a methodology for quantifying temporal clustering in our data, extending the bootstrap Kolmogorov-Smirnov tests described by Abadie (2002). We next turn our attention to the intelligence gathering mechanism.

In particular, our model suggests that the quality of information that insurgents can gather about troop and facility weakness influences the level of sophistication exhibited by the temporal patterns of combat activity. As insurgents gather enough sufficiently precise information about target vulnerabilities, their attack patterns will shift as they optimally calibrate the timing of their attacks. Our military records include previously unreleased information about rebel spy operations (surveillance of troop movement and base activity) observed by military forces. Our data suggests that insurgents were able to conduct surveillance operations in 70 of Afghanistan's 398 districts in 2006 (see Figure 6). These records also include data on security breaches, which occur when insurgents are able to effectively infiltrate the outer perimeter of targets and observe activity from within bases and outposts. Finally, consistent with these measures, we track incidents where the Taliban are able to 'turn' security recruits and use them to launch demoralizing attacks within army units. In 2006, insurgents were able to infiltrate bases in four districts and conduct insider attacks in five. These data yield a unique opportunity to test whether the ability to surveil, breach, or infiltrate targets—all of which are consistent with the capacity to acquire precise information about target weaknesses—significantly enhances the overall impact of revenue shocks. We find strong evidence that the baseline effect we observe is substantially greater in areas where the Taliban have the capability to acquire actionable intelligence about when troops, convoys, and bases are susceptible to attack.

Quantitative work on the economics of conflict has largely focused on the conditions that trigger warfare. Fearon and Laitin (2003), Collier and Hoeffler (2004), Miguel et al. (2004), and Bazzi and Blattman (2014) study the proximate causes of conflict. Other work explores incentives to use violence as a means to predate or capture economic resources (Le Billon, 2001; Ross, 2004; Hidalgo et al., 2010). Still others have focused on the link between income shocks and levels of political violence at a subnational level (Dube and Vargas, 2013b; Jia, 2014; Oliver, 2018). A meaningful gap exists, as Berman and Matanock (2015) point out, between our understanding of *when* and *how* rebels engage in armed combat. We advance this agenda by examining the relationship between economic shocks to rebel capacity and the ways in which insurgents fight.

Violence is economically and politically costly. Gould and Klor (2010) find that terror attacks harden in-group biases and cause Israelis to adopt less accommodating political positions.[3] Those affected by these attacks are also more likely to vote for right-wing parties. Similarly, Condra et al. (forthcoming) find that insurgent attacks around elections in Afghanistan are calibrated to avoid civilian casualties and substantially reduce voter turnout. Violence may be triggered by and reinforce ethnic divisions (Esteban and Ray, 2011; Esteban et al., 2012), even in institutions designed to maintain impartiality (Shayo and Zussman, 2017). Civil conflict is also economically disruptive (Abadie and Gardeazabal, 2003), even at the microlevel (Besley and Mueller, 2012). Our results help us better understand how insurgents respond to rent shocks and may enable governments to more anticipate and defend against attacks.

More generally, our model helps unpack core features of the political economy of insurgency. State capacity is central to economic theories of conflict (Besley and Persson, 2011; Gennaioli and Voth, 2015; Besley and Persson, 2010; Powell, 2013; Carter, 2015; Esteban et al., 2015). 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. Recognizing this gap, Bueno de Mesquita (2013) theorizes that the relative strength of the rebellion influences leaders' choice to adopt irregular tactics. We advance this literature by developing

[3] See also Berrebi and Klor (2008) and Getmansky and Zeitzoff (2014).

a microlevel theory of the relationship between combat tactics and resource endowments.

Technology of conflict has long been an active area of theoretical modelling (Kress, 2012). Optimal allocation of attacking and defensive resources have been studied in the Colonel Blotto-type games starting with Borel (1953). (See Golman and Page (2009), Roberson and Kvasov (2012) and Konrad and Kovenock (2009) for recent advances and Kovenock and Roberson (2012) for an excellent survey). Our baseline model is very simple: the rebels maximize the expected number of successful attacks. The advantage is its tractability: it allows to add the possibility of one-sided information-gathering and comparative statics with respect to the quality of information.

In a setting in which the maximand is the sum of harm to individual sites, Powell (2007a) demonstrates that in the unique equilibrium, the defender minimaxes the attacker regardless of whether or not the attacker can observe the defender's allocation before choosing where to launch the attack. In Powell (2007a), the defense has private information about the relative vulnerability of two sites, and allocated protection is public information, which creates the secrecy vs. vulnerability trade off. Goyal and Vigier (2014) consider a zero-sum Tullock-type conflict, in which the defense chooses a network to protect and the attacker chooses where to launch an attack. In particular, they show that a star network with all defence resources allocated to the central node is optimal in many circumstances. In Konig et al. (2017), the network is given; the data on the Second Congo War is used to estimate the impact of selective dismantling of fighting groups and weapons embargoes on conflict intensity.

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

2 Theory

In this section, we present a model that has a rebel group choosing the number of attacks and the precision of information that they use in launching them. This is a simplest possible model consistent with our empirical approach that utilizes randomization inference and the bootstrap Kolmogorov-Smirnov method developed by Abadie (2002).

2.1 Setup

Consider a rebel group that attacks the government facilities using a certain technology (e.g., mortars). The group chooses the number of attacks A, a positive integer, and the quality of information $\theta \in [0, 1]$, on which the allocation of attacks is based. The government allocates resources to defeat attacks.

We assume that there are N time slots to protect.[4] The government has resources to defend $R < N$ time slots . Formally, the government strategy is a function $g : R \to \Delta_N$ that maps each of R units of government's defense into probability distribution over N time slots; Δ_N denotes the standard $N-1$-dimensional simplex, the set of all probability distributions over N time slots.

If an attack happens occur during the period when a target is defended, it does not succeed; if an attack is during unprotected time slot, it succeeds with probability p. Since any deterministic choice of protection will result in rebels attacking outside of the time slots when targets are defended, any reasonable placement of protection should be probabilistic. After the government allocates protection, rebels gather intelligence about which time slots are vulnerable. Specifically, rebels receive noisy signals $(s_i)_{i \in N} \in \{V, D\}^N$ that are determined according to $\Pr(s_i = P|P) = \Pr(s_i = D|D) = \theta$. Then the rebels strategy is a function

[4]It is trivial to point out that attacks are time-space objects. Here, to make the problem theoretically and empirically tractable, we focus on a single dimension: the timing of rebel attacks.

$a^* : A \times \{V, D\}^N \to \Delta_N$ that maps signals about periods' vulnerability into probabilities of attacks across each of the N time slots.

The rebel group maximizes the expected number successful attacks net of the cost of attacks and information gathering. We assume that the marginal costs of attacks and information precision are $MC_A(A) = c_A A$ and $MC_I = c_I \theta$, respectively. (This assumption grossly simplifies comparative statics - see Proposition 3; the qualitative results go through for any cost function with increasing marginal costs.) Periods of low rents correspond to higher marginal costs: while we do not model alternative uses of money by rebels explicitly, this is a realistic assumption (e.g., Dube and Vargas (2013a)). The government is interested in minimizing the expected number of successful attacks.

Timing

1. Rebels choose the number of attacks A and precision of information θ.

2. The government chooses allocation of resources across N time slots.

3. Rebels receive information $(s_i)_{i \in N} \in \{D, V\}^N$ and choose attack times in A^N.

4. Pay-offs are received.

Definition 1 *An equilibrium is rebels' choice of number of attacks A^*, the quality of information θ^*, a function $a^* : A \times \{V, D\}^N \to \Delta_N$ that maps signals about all times during which attacks could occur into probabilities of attacks on each of the N time slots, and a function $g^* : R \to \Delta_N$ that maps R units of government's defense on N time slots. Given g^*, (A^*, θ^*, a^*) maximize the probability of a successful attack; given (A^*, θ^*, a^*), g^* minimizes this probability.*

2.2 The Attack Timing Game

We start backwards. Suppose that rebels have chosen the number of attacks A they launch and precision of information θ. Our first goal is to describe a unique (symmetric) equilibrium of the resulting subgame.

It is straightforward to establish that the government allocates resources into R protection intervals 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" ($s_i = V$) and $N - x$ "defended" ($s_i = D$) time slots. Let $q(x)$ denote the *ex post* probability that time i with $s_i = V$ is vulnerable. Importantly, although signals are conditionally independent, a signal about vulnerability of one period is informative about the vulnerability of other periods. Indeed, if one time slot is more likely to be vulnerable, other time slots are less likely to be vulnerable as probability of being one of $R-1$ protected time slots among $N-1$ periods is smaller than to be one of R among N. It is possible that the number of "positive" (vulnerable) signals x is not equal to A; x can be any integer between 0 and N with a non-zero probability. In the two extreme cases, $x = 0$ (there are no time slots that are more likely to be vulnerable) and $x = N$ (signals suggest that all potential times an attack could occur are equally vulnerable), there is no information to update upon. In all other cases, $1 \leq x \leq N - 1$, signals are informative:

$$P(V|s_i = V) > P(V|s_i = D).$$

The number of "vulnerable" signals is a random variable, the sum of two binomial distributions with different probabilities of success: $N - G$ vulnerable time slots produce signal V with probability θ, while G defended time slots produce signal V with probability $1 - \theta$.[5]

[5] The probability distribution of a sum of two or more binomial random variables, i.e. sums of independent Bernoulli trials, with different success probabilities is sometimes called Poisson binomial distribution (Hillion and Johnson, 2017).

Proposition 1 *There exists a unique equilibrium in the attack timing game, in which the government protects R time slots chosen randomly and uniformly across all possible combinations and rebels follow the signals that they receive. If $A \leq x = \#\{s_i = V\}$, then A attacks are distributed uniformly over x vulnerable time slots. If $A > x$, there is an optimal number of attacks $\overline{a}(x)$ against a vulnerable period of time such that* $\min\{A, x\overline{a}(x)\}$ *attacks are distributed uniformly over x vulnerable time slots. The remaining $A - \min\{A, x\overline{a}(x)\}$ attacks are distributed uniformly across $N - x$ time slots that are marked "defended".*

Critically, with time slots labeled "vulnerable" and "defended" after the informative signals are received, the rebels' optimal strategy is a function of the probability $q(x)$ that a given time period deemed vulnerable is indeed vulnerable. (With fixed N and R, the probability that the time slot labeled defended is vulnerable is a function of $q(x)$.) The intuition is as follows. Consider the rebels' choice of one attack across two time slots with probabilities of being vulnerable q_1 and q_2, respectively, with $q_1 > q_2$. If there are no attacks already planned on during these times, then an attack timed during the first period provides a higher marginal probability of success.

The fact that the vulnerable time slot has a higher probability of success does not mean that all attacks should be concentrated during times when targets are vulnerable. Suppose that there are already m attacks planned during the first time period, $m \geq 1$, and no attacks planned on the second time slot. One more attack during the first time slot results in $q_1 (1-p)^m$ of marginal probability of success as with probability $1-(1-p)^m$ one of the m attacks that are already planned succeeds. An attack that occurs during the second time window contributes $q_2 p$. This explains why, given some sufficient capacity, rebels will launch some attacks launched during the second, less likely to be vulnerable, period regardless of the initial distribution of probabilities q_1 and q_2.[6]

[6] Note that it never makes sense to "divide" attacks. The reason is that a marginal increase in x, the "share" of the second attack launched at time window 2, does not affect the marginal probability of

The term $\min\{A, xa(x)\}$ appears in the statement of Proposition 1 because it is possible that the threshold $a(x)$ is such that $xa(x)$, the desired number of attacks during vulnerable time slots, does not exceed the capacity A. For example, if $N = 5$, $A = 3$, and intelligence about defensive weaknesses suggests two specific time windows are likely to be vulnerable, the optimal strategy might call for $m(2) = 2$, i.e. two attacks should be launched during each of the two vulnerable time slots. With the capacity of launching only three attacks, the rebels would have to choose the period for the double attack randomly over the two vulnerable time windows.

2.3 The Rebel's Demand for Precise Information

The rebels' equilibrium strategy described in Proposition 1 depends on the quality of information θ. In the simple environment of Example 1, we demonstrated that a higher precision of information leads to a higher temporal concentration of attacks: more attacks are launched during time periods that intelligence gathering indicated as "vulnerable". This result extends to the general setup: with more precise information, for any number of "vulnerable" signals x, the optimal strategy requires to launch more attacks during the vulnerable time windows.

Proposition 2 *For any number of attacks A, the higher is the precision of information that rebels receive, θ, the higher is the temporal clustering (concentration) of attacks: the lower*

success of this attack. Consider the following strategy by rebels in the case when when information received differentiates between the two time periods, so the posterior probability that the first time slot is protected is q and that the second window is protected is $1 - q$ with $q > \frac{1}{2}$. Suppose that one attack is launched during time window 1, and let x denote the probability that the other attack is launched during time slot 1 as well. Then the case of "two attacks during time slot 1" corresponds to $x = 1$, and "attack during each time period" corresponds to $x = 0$, but the strategy allows for a more general distribution of attacks. The probability that at least one attack is successful is given by

$$q\left(x\left(p^2 + 2p(1-p)\right) + (1-x)p\right) + (1-q)(1-x)p,$$

which is a linear function in x. Generically, the optimal choice is either to launch two attacks during a given time window 1 (when $q > \frac{1}{2-p}$), or to attack during each of the two periods, but never to "mix".

is the expected number of unique periods attacked and the larger is the expected number of attacks (both successful and total) per time slot attacked.

The critical element of Proposition 2 is that for any number x of time slots that are considered vulnerable after the information is collected, the probability $q(x)$ that a time window marked vulnerable is indeed vulnerable is (weakly) increasing in the precision of information θ, and thus the threshold $a(x)$ is (weakly) increasing in θ for any x. As a consequence, more precise information leads to a higher concentration of attacks: more attacks are launched during a smaller number of unique periods (of the day).

In the extreme case of perfect information ($\theta = 1$), attacks are randomized over all vulnerable times during which an attack could be launched. In the opposite extreme, when the signals are not informative at al ($\theta = \frac{1}{2}$), all time windows are equally likely to be vulnerable. In this case, the optimal strategy for rebels is to launch A attacks during different times chosen randomly and uniformly across all possible choices guaranteeing that two attacks are not launched during the same time window (unless, of course, $A > N$).

2.4 The Optimal Timing of Attacks

Now that we have determined the relationship between the precision of information that rebels receive and the temporal concentration of their attacks for any fixed number of attacks, we can analyze the choice between harvesting information and launching more attacks in the main game.

In the equilibrium, the optimal choice of rebels are functions of the marginal costs of an attacks c_A and information precision c_I, the number of potential time slots for attacks N, the resources in the disposal of the government R, and the efficiency of weapons p. Naturally, when the cost of an attack for rebels increases, they cut down on both the number of attacks and the precision of information that they use. As a result, their attacks become less concentrated during particular periods of the day: when rebels are relatively weak (i.e., revenue-poor), their attacks become increasingly random w.r.t. timing. (In the extreme case of no information, the maximum entropy of attacks is achieved as rebels use the uniform distribution of attacks over all possible time slots.) This increase in entropy, which is the main concept in the statistical method developed in Abadie (2002), as a result of decrease in resources is the central empirical implication of our model which tests in the next section.

The increase in government resources has the same effect, but for a different reason. A higher R decreases the probability of rebels' success as the *ex ante* probability that each time window is protected increases. Consequently, rebels spend less on information gathering and, in equilibrium, their attacks become less temporally concentrated. Proposition 3 formally states the comparative statics results.

Proposition 3 *There exists a (subgame perfect) equilibrium of the game, in which the rebels optimally choose the number of attacks and the quality of information that they gather before the attack timing game.*

(i) A higher marginal cost of information, c_I, or a higher marginal cost of an individual attack, c_A, results in the lower optimal precision of information, θ^*. Consequently, the rebels' attacks become more randomly timed (i.e., less concentrated): a smaller share of total number of attacks are launched during the same time windows.

(ii) More resources in the government's disposal, R, and less efficient weapons (lower p) result in lower demand for information, and, therefore, lower temporal concentration of attacks.

3 Empirical Design

In this section, we discuss the setting of our investigation, review our microdata, and introduce our identification strategy.

3.1 Context

We study the relationship between combat activity and rebel capacity in Afghanistan. We focus specifically on the well-documented link between opium production and Taliban tax extraction. In the primary opium producing regions, seeds are planted in late fall and early winter. The growing season ranges from February through April, with most opium latex harvested and packaged in May and June. 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).[7] 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.

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 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

[7]For a detailed account of the industrial organization of the Taliban, see Peters (2009).

insurgency, most notably the taxation and command structures oriented around administrative districts, 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.

3.2 Conflict Microdata

Our investigation exploits newly declassified military records which catalogue combat engagements and counterinsurgent operations during Operation Enduring Freedom in Afghanistan. These data were maintained by and retrieved through proper declassification channels from the U.S. Department of Defense. 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 (Weidmann, 2016). As Weidmann (2016) points out, these tactical reports represent the most complete record of the war in Afghanistan. Additional details on data collection are discussed in Shaver and Wright (2017).

The detailed nature of our conflict microdata allows us to track insurgent activity by the hour. Although this data tracks dozens of types of violence, the majority of enemy action events are characterized as indirect fire, direct fire, and IED explosions. Indirect fire consists of mortars and other weapons that can be deployed without close contact with military forces. Direct fire attacks are primarily line-of-sight, close combat events. 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.). Subsets of these data are also studied in Callen et al. (2014), Beath et al. (2013), and Condra et al. (forthcoming), and are highly comparable to tactical data collected in Iraq Berman et al. (2011).

It is important to note that insurgents have differential control over the exact timing of each of these types of combat. Indirect fire events can be initiated at any time against stationary targets. For example, a large number of mortar fire attacks were launched from hillsides and targeted military outposts. As such, insurgents maintained a high level of control over when these evente occurred. Direct fire attacks are typically frontal assaults on convoys or forward deployed troops conducting combing operations in remote villages. Although insurgents decide if to attack, conditional on having the opportunity, the timing of these attacks is partially determined by the movement of troops, the timing of which may be determined by a non-stochastic process. Similarly, roadside bombs may be planted hours or days before they are triggered. Some of these bombs may also be detonated by unintended targets, which further reduces insurgent control over the timing of these attacks. For these reasons, we focus our main analysis exclusively on the timing of indirect fire attacks.

Our military records include information about counterinsurgent operations, including find and clear missions that neutralized emplaced IEDs, discoveries of weapon caches, such as small arms, ammunition, and bomb making materials, and provision of close air support, which was used primarily to harden mobile targets and extract coalition forces that were pinned down at a fixed location by insurgents. We supplement this information with measures of insurgent capture and detention, counterinsurgent surveillance operations, and safe house raids yielding actionable intelligence about rebel operations. This information enables to address operational factors that may confound the relationship we are trying to estimate. In particular, it might be the case that counterinsurgents strategically assign additional fighting capacity to districts with a substantial rebel presence *and* where opium production is high. Shifts in government resources, like the ability to identify and neutralize weapon caches or to use aerial bombardment to assist troops engaged by insurgents, might also influence the temporal patterns of insurgent attacks.

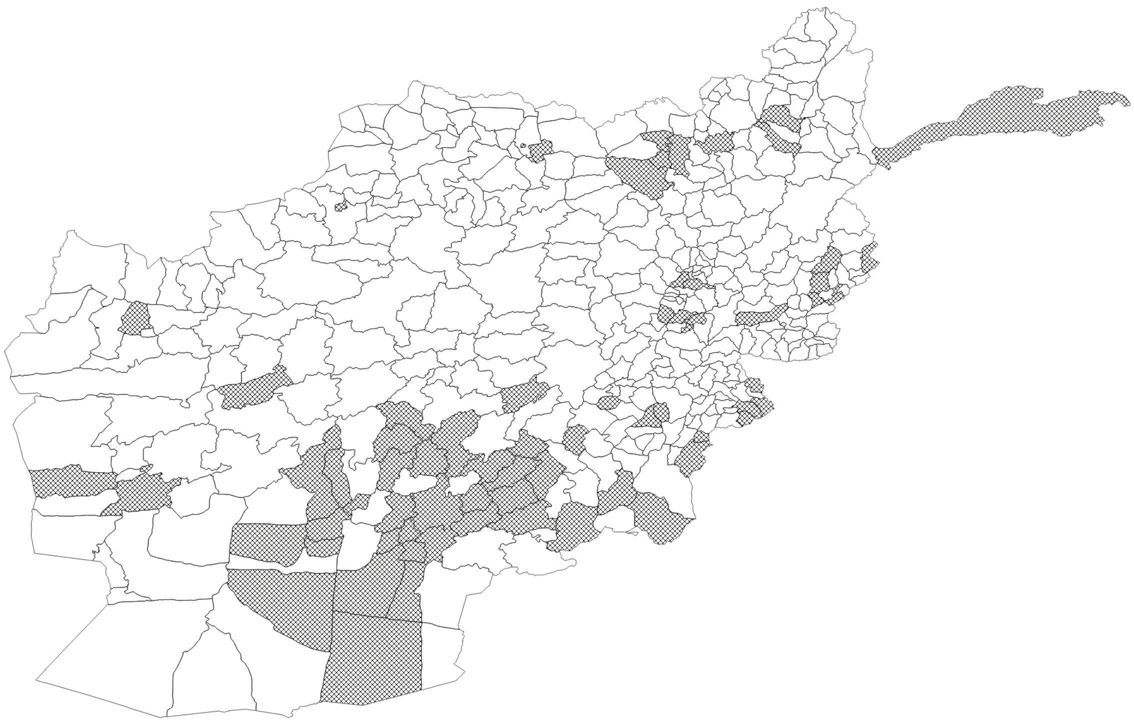
It is also possible that opium production is influenced not just be combat operations, but also more direct attempts by insurgents to coerce the local population. In particular, insurgents may use violent and non-violent tactics to intimidate civilians, such as killings of government collaborators and the posting of 'night letters' and other non-lethal shows of force. Fortunately, our military records include information about these tactics as well, enabling us to address potential concerns about residual endogeneity in estimated levels of opium production.

Another unique feature of our conflict data is information on rebel surveillance operations, military installation breaches, and insider attacks. Enemy surveillance operations are logged whenever security forces become aware of attempts by insurgents to track troop and vehicle movement and day-to-day activities on military bases. This kind of insurgent spy activity can be, and likely is, used to identify troop and infrastructure vulnerabilities. We illustrate the location of these spy operations in Figure 6. In addition to surveying force activity from outside of military compounds, insurgents can infiltrate these installations and monitor activity from within. These security breaches might also result in direct confrontations between insurgent and counterinsurgent forces. Records of insider attacks also reveal Taliban attempts to 'turn' Afghan security forces, and launch deadly attacks from within operational units. We use these incredibly unique pieces of information to more explicitly test the observable implications of our model regarding intelligence gathering and the quality (precision) of information about government vulnerabilities.

3.3 Opium Cultivation and Prices

We supplement our military records with data on opium production 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 2, panel (a)). These images are then examined for changes in vegetative signatures consistent with the volatile

Figure 1: Military records indicate the location of rebel surveillance operations conducted in Afghanistan (2006).

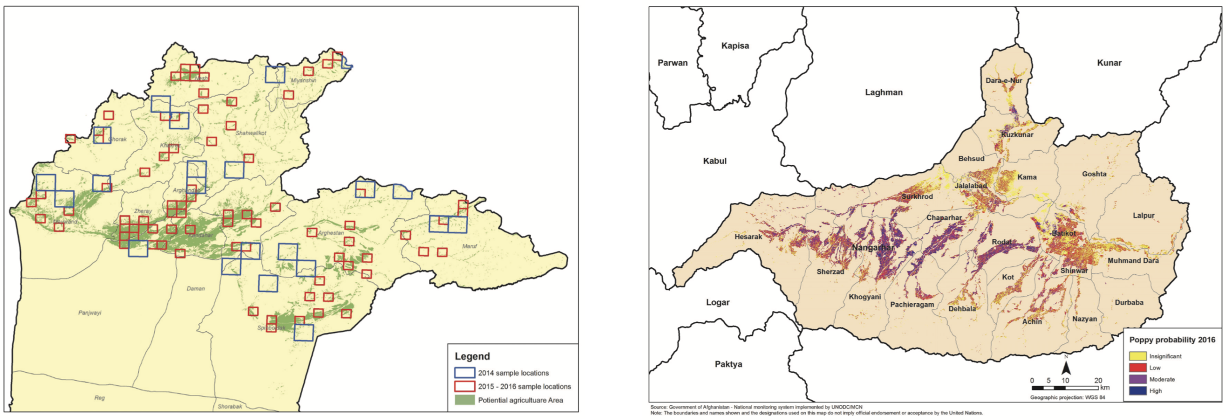


Notes: Data on insurgent spy operations drawn from SIGACTS military records. Cross hatch pattern indicates insurgents conducted at least one detected surveillance operations during 2006, the first year of our sample. District boundaries are drawn from the ESOC Afghanistan map (398 districts).

wetness of opium plants after lancing. From this sampling technique, officers estimate the spatial risk of opium production. This enables them to calculate granular estimates of opium production (see Figure 2, 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. These figures are available at the national level as well as by region.

Opium price data is compiled at national and regional levels. These prices are tracked monthly at various locations across the country via a farmer and market spot price survey system. In Figure 2, panel (a), we introduce the monthly time series for the two price-making regions, Kandahar and Nangahar. Our main specification utilizes the simple average between these prices in June, when taxes are collected (vertical lines added for clarity). In supplemental results, we study the regional price time series in Figure 2, panel (b). We rely on UNODC-Afghanistan documentation to assign districts to price zones. We extract exact prices using WebPlotDigitzer software.

Figure 2: UNODC Methodology for Estimating Annual District Drug Production.



(a) Sampling Satellite Imagery

(b) Impute Production from Imagery/Field Obs.

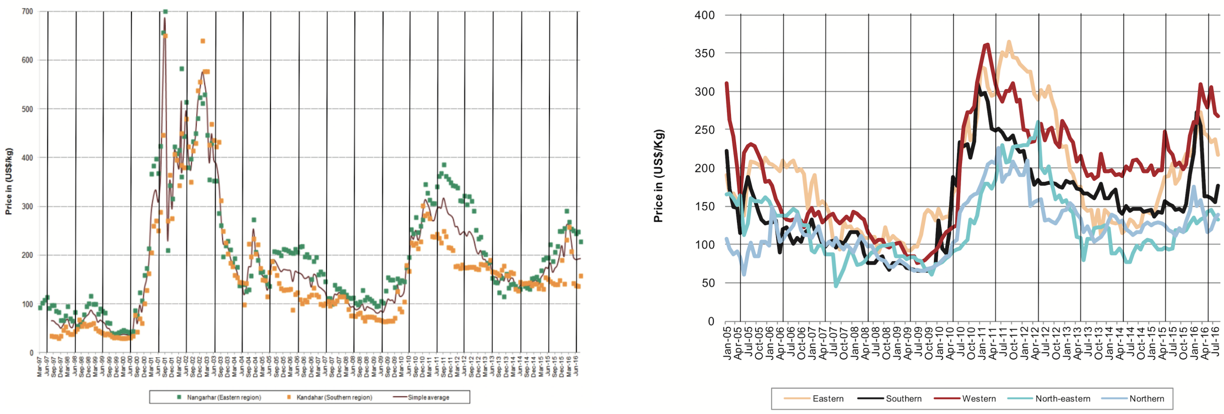
Notes: Methodological figures and details drawn from the 2016 UNODC-Afghanistan Drug Report. Panel (a) demonstrates the sampling design used when acquiring high resolution satellite imagery (location: Kandahar). Panel (b) illustrates the subsequent production estimation, which combines low and high resolution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

3.4 Detecting Random Timing of Combat

We now introduce a method for detecting the randomness of attack timing. Our approach employs randomization inference and the bootstrap Kolmogorov-Smirnov method developed by Abadie (2002). The method 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

Figure 3: Time series data on opium prices collected at national (a) and regional (b) levels.



(a) National price TS

(b) Regional prices TS

Notes: Figures on national and regional price time series drawn from the 2016 UNODC-Afghanistan Drug Report. Underlying data compiled from farmer and market spot price surveys conducted throughout the year. In Panel (a), the simple average is calculated (Nangahar, Kandahar). In Panel (b), we assign districts to regions according the UNODC documentation. Prices were precisely extracted using WebPlotDigitzer.

stored.[8]

2. Identify the sequence of district-hours during which indirect fire 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. This test is composed of four elements.

   (a) Compute the T^{KS}_{dfi} for the fitted values of the empirical and simulated distributions, where:

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

$$T_{dfi}^{KS} = \left(\frac{n_1 n_0}{n}\right)^{\frac{1}{2}} \sup_{y \in \mathbb{R}} \left|F_{1,n_1}(y) - F_{0,n_1}(y)\right|.$$

(b) Resample observations with replacement from observed and simulated distributions. Split the resampled set into two distributions and calculate $T_{dfi,b}^{KS}$. Store $T_{dfi,b}^{KS}$.

(c) Repeat prior two steps 1,000 times.

(d) Calculate and store the likelihood parameter of the tests as $\sum_{b=1}^{1000} \frac{1 T_{dfi,b}^{KS} > T_{dfi}^{KS}}{1{,}000}$.

6. Repeat steps 2 through 5 10,000 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.

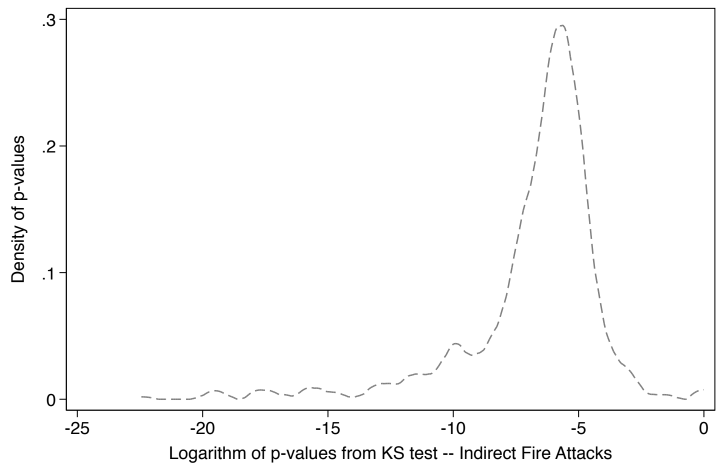
To clarify, we identify the hour of each attack within a given district-year (fighting season). We then reshuffle the hour vector and compare the empirical distribution to the randomly reshuffled vector. This process is repeated many times per district-year. The result of the technique is a single likelihood 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.[9] 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 predictable.

Estimation of our likelihood parameters using this technique requires tens of billions of simulations, so we use several supercomputers. In Figure 4, we plot the the calculated p-value (log) distribution for indirect fire attacks. Most district-year p-values above -10. This

[9]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.

distribution is characterized by a long left-side tail. This suggests that specific district-fighting seasons exhibit very clear evidence of temporal clustering (i.e., non-randomness).

Figure 4: Distribution of p-values from randomization test of combat in Afghanistan



3.5 Empirical Strategy

We study the relationship between rebel capacity and randomized combat by examining whether the within-day distribution of violence for each district's fighting season is associated with rent extraction from opium. If the results follow our expectations, rebel capacity and the likelihood parameter of our simulation test above should be negatively correlated.

To test the relationship between randomization of attack timing and insurgent capacity, we estimate the following ordinary least squares regression:

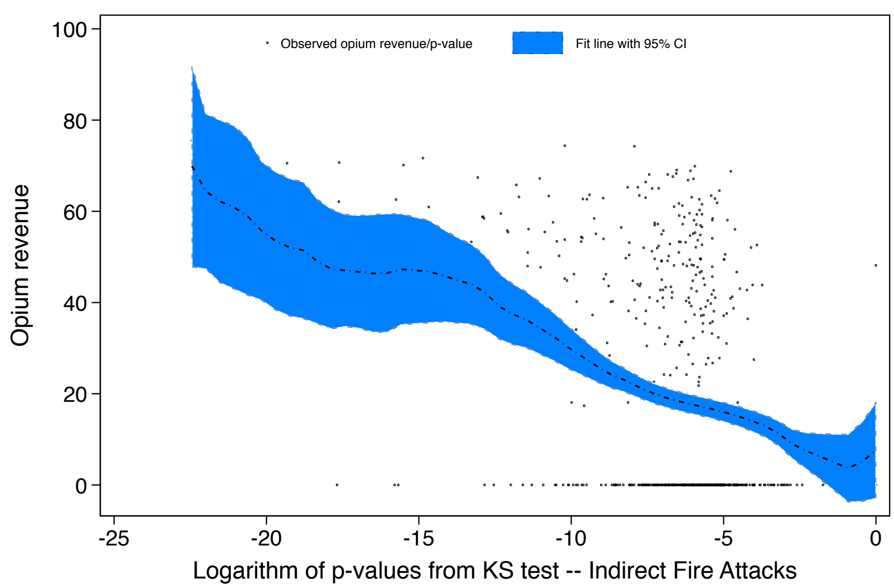
$$log(pval_{d,y}) = \alpha + \beta_1 log(production_{d,y} + 1) \times log(price_y) + \beta_2 X_y + \Lambda X^V_{d,y} + \epsilon \qquad (1)$$

Where $pval_{d,y}$ is the p-value for a given district, d, and fighting season ,y (year). X_y captures fighting season fixed effects and $X^V_{d,y}$ captures a vector of additional covariates, which we incorporate to address potential concerns about omitted variables. These covariates include the intensive margin of insurgent operations during the fighting, harvest, and planting seasons as well as supplemental measures of state capacity and insurgent intimidation. We expect the first coefficient β_1 to be negative.

4 Results

We begin by visualizing the data. In Figure 5, we plot the p-value distributions for indirect fire attacks against the corresponding opium revenue of each district-year (fighting season). Confidence regions (95%) are shaded in blue. Notice the consistently negative relationship between revenue and randomness of combat. This non-parametric correlation is consistent with our intuition that high capacity rebels produce patterns of violence that are less random and exhibit temporal clustering.

Figure 5: Bivariate relationship between opium revenue and p-value of randomization test of combat (indirect fire attacks) in Afghanistan



We next turn to our regression-based evidence. In Table 1, we estimate equation 1. Column 1 is a sparse model that demeans our bivariate correlation by fighting season. We find that a strong negative relationship between opium revenue and combat randomness, confirming our visual evidence. In Columns 2 through 4 we sequentially add controls for the intensive margin of fighting during the fighting, harvest, and planting seasons respectively. We do this to address potential concerns that our bivariate result is driven by a strong relationship between opium revenue and the intensity of violence during the fighting season (which covaries negatively with our randomness parameter). Our results are largely unaffected, although our point estimate becomes marginally more precise in Column 2. In Column 3, we attempt to rule out concerns that our estimates are substantially biased by the endogenous relationship between opium production and insurgent violence during the harvest season, which might influence subsequent conflict during the later fighting months. It might also be the case that farmers are coerced into planting opium through violence exposure. We account for this potential source of bias by including a planting season violence trend in Column 4. This baseline evidence suggests a precise, consistent link between rebel capacity and randomization of indirect fire attacks.

Table 1: Impact of rebel capacity on within-day randomization of indirect fire attacks

Opium Revenue	-0.0581***	-0.0588***	-0.0549***	-0.0549***
	(0.0137)	(0.0135)	(0.0125)	(0.0125)
Model Parameters				
Fighting Season FE	Yes	Yes	Yes	Yes
FS Trend		Yes	Yes	Yes
PS Trend			Yes	Yes
PLS Trend				Yes
Model Statistics				
No. of Observations	600	600	600	600
No. of Clusters	154	154	154	154
R^2	0.154	0.171	0.187	0.187

Notes: Outcome of interest is the (log) p-value of the randomness test. The quantity of interest is opium revenue for a given district-year. All regressions include fighting season fixed effects. Column 2-4 add controls for the intensive margin of fighting during the fighting, harvest, and planting seasons respectively. Heteroskedasticity robust standard errors clustered by district are reported in parentheses. Stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

We conduct several other baseline robustness checks in Table 2. In Column 1, we introduce a province×year fixed effect, which soaks up any residual mechanism design effects due the spatial sampling procedure utilized by the UNODC to estimate local opium production. This fixed effect also absorbs troop rotation schedules which coincide with the province-year, which includes force movement into and out of regional command posts. Although the magnitude and precision of our main effect declines marginally, the negative relationship persists even in this very demanding specification. Another potential concern one might have is that the strategy we use to account for the intensive margin of violence during the fighting season (i.e., partialling out the violence in levels) is incomplete. Another solution is to inversely weight our model along this margin. This implies that our corresponding estimate is less vulnerable to vertical (conflict) outliers. We present this result in Column 2, which is highly consistent with our baseline result.

In our baseline specification, we rely on a national time series in prices and year-by-year variation in opium yields (kilograms per hectare). Yet the interaction of weather conditions and soil suitability may lead to heterogeneous crop yields by region and year. Regional prices might also differ substantially. These two concerns represent classical error-in-variables, which we can correct with regional price and yield data compiled from UNODC records. We replicate our baseline specification in Column 3 using opium revenues calculated using regional yield rates and regional price data. It is also the case that some provinces have later harvests which occur after fighting has begun in other parts of the country. We use crop calendar maps produced by the UNODC to classify late harvest districts and exclude them from our sample in Column 4. Again, our results are highly consistent in both model specifications.

4.1 Threats to Inference

In this subsection we detail additional potential threats to inference.

Table 2: Impact of rebel capacity on within-day randomization of indirect fire attacks, robustness checks

Opium Revenue	-0.0450**	-0.0531***		-0.0543***
	(0.0220)	(0.0116)		(0.0126)
Opium Revenue (Regional)			-0.0555***	
			(0.0128)	
Model Statistics				
Fighting Season FE	Yes	Yes	Yes	Yes
FS Trend	Yes	Yes	Yes	Yes
PS Trend	Yes	Yes	Yes	Yes
PLS Trend	Yes	Yes	Yes	Yes
Prov×Year FE	Yes			
IM Weighted LS		Yes		
Regional Yield Adjust.			Yes	
Early Harvest Only				Yes
Model Statistics				
No. of Observations	600	600	600	588
No. of Clusters	154	154	154	150
R^2	0.379	0.188	0.184	0.184

Notes: Outcome of interest is the (log) p-value of the randomness test. The quantity of interest is opium revenue for a given district-year. All regressions include fighting season fixed effects as well as controls for the intensive margin of fighting during the fighting, harvest, and planting seasons respectively. Additional parameters are noted in the table footer. Heteroskedasticity robust standard errors clustered by district are reported in parentheses. Stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

It is possible that revenue from opium production may attract additional counterinsurgent investments. Consistent with our theoretical model, to defeat rebels accumulating resources, security forces may have deployed additional resources that could complicate estimation of the effect of rebel capacity on the randomness of combat. We focus on six measures of counterinsurgent capacity: close air support missions, cache discoveries, IED neutralizations, detention of insurgent forces, counterinsurgent surveillance operations, and safe house raids yielding actionable intelligence assets (salary and recruitment logs, hard drives, forensic materials, etc.). We sequentially add these covariates to equation 1. We present these results in Tables 3 and 4. For comparison, the most conservative specification from Table 1 (Column 4) is included as Column 1 in both of these tables. Notice that these measures of counterinsurgent capacity improve the explanatory power of our models, although the magnitude of the main effect is slightly attenuated.

Table 3: Impact of rebel capacity on within-day randomization of indirect fire attacks, accounting for state capacity measures (part i)

Opium Revenue	-0.0549***	-0.0336***	-0.0418***	-0.0276***
	(0.0125)	(0.00779)	(0.0106)	(0.00965)
Model Statistics				
Fighting Season FE	Yes	Yes	Yes	Yes
FS Trend	Yes	Yes	Yes	Yes
PS Trend	Yes	Yes	Yes	Yes
PLS Trend	Yes	Yes	Yes	Yes
Close Air Support		Yes		
Weapon Caches FC			Yes	
IEDs FC				Yes
Model Statistics				
Model Statistics				
No. of Observations	600	600	600	600
No. of Clusters	154	154	154	154
R^2	0.187	0.337	0.242	0.310

Notes: Outcome of interest is the (log) p-value of the randomness test. The quantity of interest is opium revenue for a given district-year. All regressions include fighting season fixed effects as well as controls for the intensive margin of fighting during the fighting, harvest, and planting seasons respectively. Additional parameters are noted in the table footer. Heteroskedasticity robust standard errors clustered by district are reported in parentheses. Stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Opium production might also be influenced by coercive tactics used by the Taliban to intimidate civilians. In general, these tactics are difficult to track. These coercive tactics represent a problematic omitted variable if they are indeed correlated with production and further correlated with the sophistication of combat tactics used during the fighting season. These are plausible concerns. To address them, we incorporate additional information from our military records which tracks attempts to intimidate the civilian population, using methods like 'night letters' and shows of non-lethal force as well as deliberate killings of government collaborators (like informants and security force recruits). Our main effect is only slightly attenuated when we account for these rebel intimidation tactics in Columns 2 and 3 of Table 5.

Table 4: Impact of rebel capacity on within-day randomization of indirect fire attacks, accounting for state capacity measures (part ii)

Opium Revenue	-0.0549***	-0.0454***	-0.0548***	-0.0484***
	(0.0125)	(0.0104)	(0.0123)	(0.0113)
Model Statistics				
Fighting Season FE	Yes	Yes	Yes	Yes
FS Trend	Yes	Yes	Yes	Yes
PS Trend	Yes	Yes	Yes	Yes
PLS Trend	Yes	Yes	Yes	Yes
Insurgent Detention		Yes		
COIN spies			Yes	
Safe house raids				Yes
Model Statistics				
No. of Observations	600	600	600	600
No. of Clusters	154	154	154	154
R^2	0.187	0.249	0.187	0.217

Notes: Outcome of interest is the (log) p-value of the randomness test. The quantity of interest is opium revenue for a given district-year. All regressions include fighting season fixed effects as well as controls for the intensive margin of fighting during the fighting, harvest, and planting seasons respectively. Additional parameters are noted in the table footer. Heteroskedasticity robust standard errors clustered by district are reported in parentheses. Stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

4.2 Mechanisms: Heterogeneous Effects of Intelligence Gathering

Having addressed numerous potential threats to inference, we now turn our attention to a more direct test of our model's empirical implications. Our results could be working through at least two plausible mechanisms. First, revenue from the opium trade could give field commanders more flexibility to recruit and arm fighters. This would enable them to engage in combat operations where the timing of attacks is less random and, therefore, easier for state rivals to anticipate and engage in strategic adjustment. As such, battlefield losses are likely. These losses are more easily absorbed if a given rebel division has more, better armed combatants. Second, increasing capacity could make it easier for insurgents to deploy spies or buy information about troop movement patterns from civilians. Responding strategically to intelligence reports about time windows within which troops and bases are vulnerable could lead to temporal clustering. Importantly, these are not rival mechanisms. They can

Table 5: Impact of rebel capacity on within-day randomization of indirect fire attacks, accounting for rebel intimidation tactics

Opium Revenue	-0.0549***	-0.0492***	-0.0484***
	(0.0125)	(0.0121)	(0.0104)
Model Parameters			
Fighting Season FE	Yes	Yes	Yes
FS Trend	Yes	Yes	Yes
PS Trend	Yes	Yes	Yes
PLS Trend	Yes	Yes	Yes
Intimidation		Yes	
Collaborator Killings			Yes
Model Statistics			
No. of Observations	600	600	600
No. of Clusters	154	154	154
R^2	0.187	0.204	0.227

Notes: Outcome of interest is the (log) p-value of the randomness test. The quantity of interest is opium revenue for a given district-year. All regressions include fighting season fixed effects as well as controls for the intensive margin of fighting during the fighting, harvest, and planting seasons respectively. Additional parameters are noted in the table footer. Heteroskedasticity robust standard errors clustered by district are reported in parentheses. Stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

operate simultaneously.

Our data provides us with a unique opportunity to investigate the second mechanism more precisely. In particular, we expect that conditions that make intelligence gathering easier (i.e., reduced frictions) should enhance the negative effect of opium revenue on randomness of combat. That is, in places where rebels have the ability to gather information about troop and base vulnerabilities, temporal clustering should be even more responsive to rents extracted from the opium trade.

We begin with the simplest test of this mechanism. We gather administrative data on the distribution of ethnic groups across districts. We identify districts where 95% or more of population settlements are classified as Pashto speaking. Pashtuns are Taliban coethnics and form the primary base of civilian support for the insurgency. In principle, we expect it will be easier for insurgents to acquire information about government activity in districts dominated by their coethnics. We present these results in Table 6, Column 2. Notice that the additive effect of opium revenue in coethnic zones is much larger (roughly 80%) than non-coethnic districts. This is evidence consistent with our story, but relying on coethnicity as a test of our argument may conflate intelligence frictions with a range of other factors as well.

We overcome this concern by turning to more sophisticated tests of our argument. In particular, our military records include previously unreleased information about rebel surveillance of troop movement and base activity as well as data on security breaches, which occur when insurgents are able to effectively infiltrate the outer perimeter of targets and observe activity from within bases and outposts. These yield another potential test of our information mechanism by tracking incidents where the Taliban are able to 'turn' security recruits and use them to launch attacks from within army units. In each of these cases, we are able to employ much more direct evidence of the ability of the insurgency to gather precise information about target defenses. We test our mechanism using these data in Columns 3 through 5 in Table 6. In Column 3, we interact our measure of the spy network operations with revenue.[10] In Columns 4 and 5, we repeat this specification with measures of infiltration and insider attacks. We find strong evidence that our main effect is enhanced in districts where the insurgents have a demonstrated capability to conduct surveillance, infiltrate security installations, and launch insider attacks. These findings yield evidence consistent with the mechanism implied by our theoretical argument and suggest that intelligence gathering may be a primary pathway through which shocks to rebel capacity influence the timing of violent attacks.

[10]To avoid potential concerns about the endogeneity of intelligence gathering, we identify the cross section of districts with these characteristics using only the first year of our sample, 2006.

Table 6: Heterogeneous effects of rebel capacity on within-day randomization of indirect fire attacks with respect to potential intelligence gathering

Opium Revenue	-0.0549***	-0.0146*	-0.0194***	-0.0537***	-0.0433***
	(0.0125)	(0.00766)	(0.00664)	(0.0131)	(0.00920)
Coethnicity		0.378			
		(0.352)			
Coethnicity × Revenue		-0.0552***			
		(0.0168)			
Surveillance			0.953**		
			(0.400)		
Surveillance × Revenue			-0.0678***		
			(0.0191)		
Infiltration				0.761**	
				(0.307)	
Infiltration × Revenue				-0.0308*	
				(0.0161)	
Insiders					-1.861
					(1.511)
Insiders × Revenue					-0.0909***
					(0.0344)
Model Parameters					
Fighting Season FE	Yes	Yes	Yes	Yes	Yes
FS Trend	Yes	Yes	Yes	Yes	Yes
PS Trend	Yes	Yes	Yes	Yes	Yes
PLS Trend	Yes	Yes	Yes	Yes	Yes
Coethnicity		Yes			
Rebel Spies			Yes		
Rebel Infiltration				Yes	
Insider Attacks					Yes
Model Statistics					
No. of Observations	600	600	600	600	600
No. of Clusters	154	154	154	154	154
R^2	0.187	0.218	0.233	0.188	0.278

Notes: Outcome of interest is the (log) p-value of the randomness test. The quantity of interest is opium revenue for a given district-year. All regressions include fighting season fixed effects as well as controls for the intensive margin of fighting during the fighting, harvest, and planting seasons respectively. Additional parameters are noted in the table footer. Heteroskedasticity robust standard errors clustered by district are reported in parentheses. Stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

5 Conclusion

Rebel tactics are an overlooked feature of internal warfare. Understanding how these conflicts are fought and the strategic responses of armed actors to revenue shocks is on equal footing with thoroughly examined questions about the causes of civil war and the factors that influence when hostilities end.

We argue that shocks to rebel capacity influence the timing of attacks. To clarify our argument, we develop a model of combat during an irregular insurgency. Rebels optimize when they conduct attacks after observing imperfect signals of government defensive maneuvers. As insurgents accumulate resources through taxation, their budget constraint is relaxed and they can recruit more fighters, acquire more armaments, and gather more intelligence about the vulnerability of troops and bases. The labor supply of fighters, coupled with surplus capital, allows rebels to conduct more attacks. Infiltrating military installations and conducting surveillance of troop movement enhances the precision (quality) of information rebel leaders have about specific times when attacks will yield the highest probability of success. Our model implies revenue shocks will lead to clustering in the temporal distribution of violence. Stated differently, the timing of attack patterns will increasingly deviate from randomness. Our model also suggests that these tactical shifts will be greatest when armed actors have the institutional capacity to pin-point when such attacks are least likely to be neutralized by government defenses.

We test the empirical implications of our theoretical model using data collected during Operational Enduring Freedom in Afghanistan. The temporal precision of our military records enable us to develop a sophisticated methodology for differentiating the within-day timing of rebel operations from randomized combat. This method produces a likelihood parameter that quantifies the degree of temporal clustering present in the allocation of insurgent attacks at the district level, broken down by fighting season. We couple this novel approach with high resolution estimates of opium production and market prices. We leverage the industrial organization of the Taliban, including their highly institutionalized taxation system, to estimate the impact of local revenue from the drug trade on combat tactics in the subsequent fighting season.

We find consistent evidence that positive shocks to rebel capacity decrease randomness in the timing of violent attacks. Consistent with our model, indirect fire attacks (over which rebels maintain unilateral control over timing) become concentrated during particular time windows following opium tax windfalls. In other words, insurgents begin to concentrate their combat operations during specific periods of the day. This finding survives a battery of robustness checks, including accounting for trends in violence during the fighting, harvest, and planting seasons. These results also hold when we introduce a number of alternative measures of state capacity, including close air support missions, bomb neutralization, counterinsurgent surveillance activity, and safe house raids. The incredibly rich nature of our conflict microdata also allow us to rule out other potential identification concerns, including rarely documented coercive tactics used by insurgents which may influence opium production.

Our data also yields a unique opportunity to assess the mechanism suggested by our model: intelligence gathering. Temporal clustering may occur because rebels use their resources to improve the quality and precision of their information about target vulnerability. This type of mechanism is largely unobserved and difficult to disentangle from alternative explanations. Our military records, however, include detailed information about where rebels were observed conducting surveillance operations as well as instances where insurgents were able to infiltrate government installations and conduct insider attacks. The heterogeneous effects we estimate produce evidence consistent with our theoretical model. The capacity to gather intelligence substantially enhances the baseline effect of revenue shocks on combat operations.

These findings reveal how rebel tactics are influenced by the ability to extract taxes from the local population. As revenue increases, insurgents shift their combat operations in a manner that prior theoretical accounts overlooked. The timing of their attacks becomes less plausibly random and more easily predictable. This type of tactical shift could make it easier for government forces to thwart rebel attacks. But our results suggest that positive shocks to rebel capacity decrease the randomness of attacks the greatest under conditions when rebels can acquire information about government defensive maneuvers, potentially anticipate countermeasures, and calibrate the timing of their attacks accordingly.

Our results further unpack the political economy of conflict, with a focus on the industrial organization of insurgency. Prior work provides compelling evidence that insurgents respond strategically to local economic shocks (Dube and Vargas, 2013b; Berman et al., 2017; Oliver, 2018), form alliances during war (Konig et al., 2017), and calibrate their use of violence against civilian populations (Condra and Shapiro, 2012; Condra et al., forthcoming). Our evidence highlights how sophisticated institutions commonly associated with states and government forces—structured tax collection schemes, combat coordination, and surveillance operations—influence the timing of insurgent violence.

References

Alberto Abadie. Bootstrap tests for distributional treatment effects in instrumental variable models. *Journal of the American Statistical Association*, 97(457):284–292, 2002.

Alberto Abadie and Javier Gardeazabal. The economic costs of conflict: A case study of the basque country. *American Economic Review*, 93(1):113–132, March 2003.

Samuel Bazzi and Christopher Blattman. Economic shocks and conflict: evidence from commodity prices. *American Economic Journal: Macroeconomics*, 6(4):1–38, 2014.

Andrew Beath, Fotini Christia, and Ruben Enikolopov. Winning hearts and minds through development: Evidence from a field experiment in afghanistan. In *APSA 2013 Annual Meeting Paper*, 2013.

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

Eli Berman, Michael Callen, Joseph H. Felter, and Jacob N. Shapiro. Do working men rebel? insurgency and unemployment in afghanistan, iraq, and the philippines. *Journal of Conflict Resolution*, 55:496–528, 2011.

Nicolas Berman, Mathieu Couttenier, Dominic Rohner, and Mathias Thoenig. This mine is mine! how minerals fuel conflicts in africa. *American Economic Review*, 107(6):1564–1610, June 2017.

Claude Berrebi and Esteban F Klor. Are voters sensitive to terrorism? direct evidence from the israeli electorate. *American Political Science Review*, 102(3):279–301, 2008.

Timothy Besley and Hannes Mueller. Estimating the peace dividend: The impact of violence on house prices in northern ireland. *American Economic Review*, 102(2):810–33, April 2012.

Timothy Besley and Torsten Persson. State capacity, conflict and development. *Econometrica*, 78:1–34, 2010.

Timothy Besley and Torsten Persson. The Logic of Political Violence. *The Quarterly Journal of Economics*, 126(3):1411–1445, 2011. URL `https://ideas.repec.org/a/oup/qjecon/v126y2011i3p1411-1445.html`.

Emil Borel. 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, 1953.

Ethan Bueno de Mesquita. Rebel tactics. *Journal of Political Economy*, 121(2):323–357, 2013.

Michael Callen, Mohammad Isaqzadeh, James D. Long, and Charles Sprenger. Violence and risk preference: Experimental evidence from afghanistan. *American Economic Review*, 104(1):123–148, 2014.

David Carter. When terrorism is evidence of state success: securing the state against territorial groups. *Oxford Economic Papers*, 2014 (advance access):1–17, 2015.

Paul Collier and Anke Hoeffler. Greed and grievance in civil war. *Oxford Economic Papers*, 56(563-595), 2004.

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

Luke N. Condra, James D. Long, Andrew C. Shaver, and Austin L. Wright. The logic of insurgent electoral violence. *American Economic Review*, forthcoming.

Oeindrila Dube and Juan F Vargas. Commodity price shocks and civil conflict: Evidence from colombia. *The Review of Economic Studies*, 80(4):1384–1421, 2013a.

Oeindrilla Dube and Juan Vargas. Commodity price shocks and civil conflict: evidence from colombia. *Review of Economic Studies*, 80:1384–1421, 2013b.

Joan Esteban and Debraj Ray. Linking conflict to inequality and polarization. *American Economic Review*, 101(4):1345–74, June 2011.

Joan Esteban, Laura Mayoral, and Debraj Ray. Ethnicity and conflict: An empirical study. *American Economic Review*, 102(4):1310–42, June 2012.

Joan Esteban, Massimo Morelli, and Dominic Rohner. Strategic mass killings. *Journal of Political Economy*, 123(5):1087–1132, 2015.

James Fearon and David Laitin. Ethnicity, insurgency and civil war. *American Political Science Review*, 97(1):75–90, 2003.

David Galula. *Counterinsurgency Warfare: Theory and Practice*. Praeger, New York, 1964.

Nicola Gennaioli and Hans-Joachim Voth. State capacity and military conflict. *The Review of Economic Studies*, 82(4):1409–1448, 2015.

Anna Getmansky and Thomas Zeitzoff. Terrorism and voting: The effect of rocket threat on voting in israeli elections. *American Political Science Review*, 108(3):588–604, 2014.

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

Eric D. Gould and Esteban F. Klor. Does terrorism work?*. *The Quarterly Journal of Economics*, 125(4):1459–1510, 2010.

Sanjeev Goyal and Adrien Vigier. Attack, Defence, and Contagion in Networks. *Review of Economic Studies*, 81(4):1518–1542, 2014. URL `https://ideas.repec.org/a/oup/restud/v81y2014i4p1518-1542.html`.

Daniel Hidalgo, Suresh Naidu, Simeon Nichter, and Neal Richardson. Economic determinants of land invasions. *Review of Economics and Statistics*, 92(3):505–523, 2010.

Erwan Hillion and Oliver Johnson. A proof of the sheppolkin entropy concavity conjecture. *Bernoulli*, 23(4B):3638–3649, 11 2017. doi: 10.3150/16-BEJ860. URL `https://doi.org/10.3150/16-BEJ860`.

Ruixue Jia. Weather shocks, sweet potatoes and peasant revolts in historical china. *The Economic Journal*, pages 92–118, March 2014.

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

Michael D. Konig, Dominic Rohner, Mathias Thoenig, and Fabrizio Zilibotti. Networks in Conflict: Theory and Evidence From the Great War of Africa. *Econometrica*, 85:1093–1132, July 2017. URL `https://ideas.repec.org/a/wly/emetrp/v85y2017ip1093-1132.html`.

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

Dan Kovenock and Brian Roberson. 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. 2012.

Moshe Kress. Modeling armed conflicts. *Science*, 336(6083):865–869, 2012. ISSN 0036-8075. doi: 10.1126/science.1217724. URL `http://science.sciencemag.org/content/336/6083/865`.

Philippe Le Billon. The political ecology of war: Natural resources and armed conflict. *Political Geography*, 20:561–584, 2001.

Edward Miguel, Shanker Satyanath, and Ernest Sergenti. Economic shocks and civil conflict: an instrumental variables approach. *Journal of Political Economy*, 112(4):725–753, 2004.

Vanden Eynde Oliver. Targets of violence: Evidence from india's naxalite conflict. *The Economic Journal*, 128(609):887–916, 2018.

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

Robert Powell. Allocating defensive resources with private information about vulnerability. *American Political Science Review*, 101(4):799–809, 2007a.

Robert Powell. Defending against terrorist attacks with limited resources. *American Political Science Review*, 101(3):527–541, 2007b.

Robert Powell. Monopolizing violence and consolidating power. *Quarterly Journal of Economics*, pages 807–859, 2013.

Brian Roberson and Dmitriy Kvasov. The non-constant-sum Colonel Blotto game. *Economic Theory*, 51(2):397–433, 2012.

Michael Ross. What do we know about natural resources and civil war? *Journal of Peace Research*, 41:337–356, 2004.

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

Andrew Shaver and Austin Wright. Data on combatant activity during afghanistan war advance scientific investigation of insurgency. 2017.

Moses Shayo and Asaf Zussman. Conflict and the persistence of ethnic bias. *American Economic Journal: Applied Economics*, 9(4):137–65, October 2017.

Robert Thompson. *Defeating Communist Insurgency: The Lessons of Malaya and Vietnam.* Praeger, New York, 1966.

Nils B Weidmann. A closer look at reporting bias in conflict event data. *American Journal of Political Science*, 60(1):206–218, 2016.

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

Reed M. Wood. 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, 2014.

Table 7: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Outcome of Interest					
Indirect fire, likelihood parameter	-7.025	3.918	-39.686	0	600
Rebel Capacity					
Opium revenue	20.506	25.403	0	75.518	600
Opium revenue, regional yield/prices	20.017	24.7	0	73.587	600
Trends in Violence					
Indirect fire trend, fighting season	19.187	21.153	5	253	600
Indirect fire trend, growing/harvest season	10.788	15.194	0	161	600
Indirect fire trend, planting season	7.663	10.925	0	109	600

Notes: summary statistics are calculated for the sample studied in the main estimating equation.

Theoretical Appendix

Proof of Proposition 1

For event S, the complement is denoted S'.

For each time slot, we say that H denotes the situation when protection is there, and D signals that the time window is "defended"; $a = P(D|H)$, $b = P(V|H')$. Finally, let C explosion (attack is successful) during a period with an attack. We assume that $P(C|H') = p$ and $P(C|H) = 0$.

The other parameters are: N is the number of time slots, G is the number of defenses allocated by the government, and A is the number of attacks.

A standard interpretation for a and b as $a = P(\text{positive test—disease}) = \text{sensitivity}$; and $b = P(\text{negative test—no disease}) = \text{specificity}$, two important characteristics of any statistical test.

Suppose that every time slot is protected with probability t; it is specified by distribution of G, $G < N$ defenses among N time slots. Rebels have A attacks, test every time slot, and thus obtain a signal "defended" (D) or "vulnerable" (V) in every time slot with probabilities

$$\begin{aligned} P(V) &= P(H)P(V|H) + P(H')P(V|H') = t(1-a) + (1-t)b, \\ P(D) &= P(H)P(D|H) + P(H')P(D|H') = ta + (1-t)(1-b). \end{aligned} \tag{A1}$$

The optimal strategy for the government is to allocate G defenses uniformly over N time windows. Then for each window $t(G) = P(\text{time slot is protected}) = \frac{G}{N}$.

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 time slots. If $A > x$ then there is threshold value how many more put in the vulnerable time windows; then, they start to put into the windows 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 time slots. In what follows, we will assume that $0 < x < N$.

The total number of windows that are labeled vulnerable is $N \equiv N_{N,G} = N_1 + N_2$, where N_1 is number of slots that are vulnerable among the protected time slots and N_2 in unprotected, i.e. N_i are independent binomial random variables with parameters $(G, 1-a)$ and $(N-G, b)$. Then $EN = G(1-a) + (N-G)b$. We denote the event $N(x) = (N = x)$, and the p.d.f. of $N(x)$ as $g_{N,G}(x) = g_{N,G}(x|a,b))$, $0 \le x \le N$. This p.d.f. can be calculated by the standard discrete convolution formula

$$g_{N,G}(x) = \sum_{i=0}^{\min(x,G)} p_1(i)p_2(x-i).$$

The critical ratio is defined by the following equation:

$$r_{N,G}(x) = r_{N,G}(x|a,b) = \frac{P(C|VN(x))}{P(C|DN(x))}.$$

Our aim is to establish that $r_{N,G}(x) \geq 1$.

Lemma A1 *The following formulae are true for any x, $0 < x < N$:*

$$\begin{aligned} P(CVN(x)) &= \frac{G}{N} p b g_{N-1,G}(x-1), & \text{(A2)} \\ P(VN(x)) &= g_{N,G}(x)\frac{x}{N}, & \text{(A3)} \\ P(CDN(x)) &= \frac{G}{N} p(1-b) g_{N-1,G}(x), & \text{(A4)} \\ P(DN(x)) &= g_{N,G}(x)\frac{N-x}{N}. & \text{(A5)} \end{aligned}$$

Proof. First, observe that

$$P(VN(x)) = P(N(x))P(V|N(x)) = g_{N,G}(x)\frac{x}{N}$$

as $P(V|N(x)) = \frac{x}{N}$, the probability for one slot to be labeled vulnerable among x to be in a particular time window.

Similarly,

$$P(DN(x)) = P(N(x))P(D|N(x)) = g_{N,G}(x)\frac{N-x}{N}.$$

Though events C and V (or D) are not independent, they are independent given H or H'. Then, using total probability formula and $P(C|H) = 0$, we have

$$\begin{aligned} P(CVN(x)) &= P(H)P(C|H)P(VN(x)|H) + P(H')P(C|H')P(VN(x)|H') \\ &= P(H')P(C|H')P(VN(x)|H'). \end{aligned}$$

Similarly,

$$\begin{aligned} P(CDN(x)) &= P(H')P(C|H')P(DN(x)|H'), \\ P(VN(x)|H') &= P(V|H')P(N(x)|VH'), \\ P(DN(x)|H') &= P(D|H')P(N(x)|DH'). \end{aligned}$$

Now, we can show that

$$\begin{aligned} P(N(x))|H'V) &= g_{N-1,G}(x-1), \\ P(N(x))|H'D) &= g_{N-1,G}(x). \end{aligned}$$

Indeed, if the total number of time slots marked "vulnerable" is x, and if a particular window has no defense, H', and is marked "vulnerable", V, then in remaining $N-1$ time slots there are G defenses and $x-1$ time slots that are marked vulnerable. If the total number of time slots marked vulnerable is x, and a particular slot has no defense, H', and produced mark D, then in remaining $N-1$ windows there are G defended windows and x slots marked vulnerable. We also have $P(H') = \frac{G}{N}$, $P(C|H') = p$, $P(V|H') = b$, and $P(D|H') = 1-b$. ∎

Now the following lemma that describes $r_{N,G}(x)$ is a straightforward corollary.

Lemma A2 *The ratio is given by formula*

$$r_{N,G}(x) = \frac{P(C|VN(x))}{P(C|DN(x))} = \frac{b}{(1-b)}\frac{(N-x)}{x}\frac{g_{N-1,G}(x-1)}{g_{N-1,G}(x)}.$$

Lemma A3 *Suppose that the signals are informative, i.e.* $a, b \geq 1/2$ *and* $a+b>1$. *For any* N, G, $0<x<N$, *functions* $r_{N,G}(x) > 1$ *and are monotonically increasing in* x.

Proof. Let $D_{N,G}$ be a sum of N Bernoulli random variables, G have parameter $1-a$, and $N-G$ have parameter $b > 1-a$, $0 \leq G \leq N$. Let

$$\begin{aligned} g_{N,G}(x) &= P(D_{N,G}=x),\ 0 \leq x \leq N; \\ f_{N,G}(x) &= \frac{g_{N,G}(x-1)}{g_{N,G}(x)},\ 1 \leq x \leq N; \end{aligned}$$

$$r_{N,G}(x) = \frac{N+1-x}{x}\frac{b}{1-b}f_{N,G}(x), 1 \leq x \leq N.$$

We assume also that $f_{N,G}(0) = r_{N,G}(0) = 0$.

Define $c = \frac{b}{1-b}\frac{a}{1-a}$, and observe that $b+a>1$ implies $c>1$.

It is easy to check that

$$r_{N,0}(x) = 1,\ r_{N,N}(x) = c,\ \text{for } 1 \leq x \leq N\,;$$

$$r_{N,G}(N) = \frac{Gc+N-G}{N} = 1 + \frac{G}{N}(c-1) \text{ for } 0 \leq G \leq N\,. \tag{A6}$$

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

$$\begin{aligned} f_{N,G}(x) &= \frac{g_{N,G}(x-1)}{g_{N,G}(x)} = \frac{(1-b)g_{N-1,G}(x-1) + bg_{N-1,G}(x-2)}{(1-b)g_{N-1,G}(x) + bg_{N-1,G}(x-1)} \\ &= \frac{1-b+bf_{N-1,G}(x-1)}{\dfrac{1-b}{f_{N-1,G}(x)}+b} \text{ for } 1 \leq G,\ x \leq N-1\,. \end{aligned} \tag{A7}$$

This implies that

$$\begin{aligned} r_{N,G}(x) &= \frac{N-x+1}{x}\frac{b}{1-b}\frac{1-b+b\dfrac{x-1}{N-x+1}\dfrac{1-b}{b}r_{N-1,G}(x-1)}{\dfrac{1-b}{r_{N-1,G}(x)}\dfrac{N-x}{x}\dfrac{1-b}{b}+b} \\ &= \frac{N-x+1+(x-1)r_{N-1,G}(x-1)}{\dfrac{N-x}{r_{N-1,G}(x)}+x} \text{ for } 1 \leq G,\ x \leq N-1\,. \end{aligned} \tag{A8}$$

Now, let us prove the following claim. *For any $N \geq 1$, $1 \leq G \leq N$, $x \leq N$, function $r_{N,G}(x)$ as a function of a and b, depends only on c. When c grows from 1 to ∞, it is strictly increasing from 1.*

The proof of the above claim follows from (A6) and induction by N, based on (A8). Indeed, (A6) implies that the induction statement holds for $N = 1$. By induction, the numerator in (A8) is strictly increasing, and the denominator is strictly decreasing, and hence the right side in (A8) is strictly increasing and depends only on c.

Now we can show that for fixed $N \geq 2$, $1 \leq G \leq N-1$, $c > 1$, function $r_{N,G}(x)$ is strictly increasing in x.

Again, we use induction by N. Let $r(x) = r_{N-1,G}(x)$, $H(x) = N - x + xr(x)$. Then, by (A8),

$$r_{N,G}(x) = r(x)\frac{H(x-1)}{H(x)}, \quad r_{N,G}(x+1) - r_{N,G}(x) = \frac{C(x)}{H(x+1)H(x)},$$

where

$$\begin{aligned}
C(x) &= r(x+1)H^2(x) - r(x)H(x+1)H(x-1) \\
&= r(x+1)\left[(N-x)^2 + 2x(N-x)r(x) + x^2r^2(x)\right] \\
&= -r(x)[N-x-1+(x+1)r(x+1)]\,[N-x+1+(x-1)r(x-1)] \\
&= r(x+1)\left[(N-x)^2 + 2x(N-x)r(x) + x^2r^2(x)\right].
\end{aligned}$$

We can check, using (A7) and (A8), that

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

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 hence by induction (**??**), it follows that

$$C(x) > r(x) + r^2(x)r(x+1) - r(x)r(x+1) - r^2(x) = r(x)(r(x+1)-1)(r(x)-1) > 0$$

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

To prove the Lemma, it remains to show that $r_{N-G,G}(2) - r_{N-G,G}(1) > 0$ and $r_{N-G,G}(N) - r_{N-G,G}(N-1) > 0$. Using the total probability formula,

$$\begin{aligned}
r_{N,G}(N-1) = &\ \frac{2}{N-1}\,\frac{b}{1-b}\,\frac{P(D_{NG} = N-2)}{P(D_{NG} = N-1)} \\
= &\ \frac{1}{N-1}\,\frac{(N-G)(N-G-1) + 2(N-G)Gc + G(G-1)c^2}{N-G+cG}.
\end{aligned} \tag{A9}$$

Then, using (A7), we obtain

$$r_{N,G}(N) - r_{N,G}(N-1) = \frac{(N-G)G(c-1)^2}{N(N-1)(N-G+cG)} > 0\,.$$

Similarly,

$$r_{N,G}(2) = \frac{(N-1)(c(N-G)+G)}{c^2(N-G)(N-G-1)+2(N-G)Gc+G(G-1)}\,,$$

$$r_{N,G}(1) = \frac{Nc}{G+c(1-G)}\,,$$

$$r_{N,G}(2) - r_{N,G}(1) = \frac{(N-G)G(c-1)^2}{[c^2(N-G)(N-G-1)+2(N-G)Gc+G(G-1)](G+c(1-G))}\,.$$

■

Proof of Proposition 2

In our main setup $a = b = \theta > \frac{1}{2}$, which yields that

$$c(\theta) = \frac{b}{1-b}\frac{a}{1-a} = \frac{\theta^2}{(1-\theta)^2} > 1.$$

Also, $c(\theta)$ is strictly increasing in θ. Thus, Lemma applies: the critical ratio

$$r_{N,G}(x) = r_{N,G}(x|a=\theta, b=\theta))$$

is strictly increasing in θ. This in turns yields that the threshold $\overline{a}(x)$ that defines the optimal number of mortar attacks that should be allocated to vulnerable time slots is (weakly) increasing in θ, i.e. more attacks are concentrated on the vulnerable time slots. ■

Proof of Proposition 3

(i) Let $\varphi = \varphi(\theta)$ be the probability of success for the rebels as a function of the precision of information θ. Then $\varphi(\theta) - \frac{1}{2}c_I\theta^2$ satisfies the single-crossing property (in fact, increasing differences property) in $(-c_I, \theta)$. Now, we can use Theorem 4 in Milgrom and Shannon (1994) to show that $\arg\max_\theta \left\{\varphi(\theta) - \frac{1}{2}c_I\theta^2\right\}$ is a monotone non-increasing function of c_I. By Lemma A3, a fall in θ^* implies that attacks become less concentrated as the marginal cost of information increases.

Similarly to (i), $\varphi(\theta) - \frac{1}{2}c_A A^2$ satisfies the single-crossing property (in fact, increasing differences property) in $(-c_A, A)$, which implies that $\arg\max_\theta \left\{\varphi(\theta) - \frac{1}{2}c_A A^2\right\}$ is a monotone non-increasing function of c_A.

(ii) Consider φ as a function of both θ and R, and observe that φ is increasing in θ and decreasing in R, which implies that $\varphi(\theta, R)$ satisfies the single-crossing property in $(-R, \theta)$. (We ignored the cost terms, which are independent of R.) Again, Theorem 4 in Milgrom and Shannon (1994) shows that $\arg\max_\theta \varphi(\theta, R)$ is monotone non-increasing in R.

By Lemma A3, a fall in θ^* implies that attacks become less concentrated as the marginal cost of information increases.

Now, let us consider φ as a function of both θ and p, and observe that φ is increasing in both θ and p, which implies that $\varphi(\theta, p)$ satisfies the single-crossing property in (θ, p). (Again, ignoring the cost terms, which are independent of the success probability p.) This is sufficient to conclude that $\arg\max_\theta \varphi(\theta, p)$ is monotone non-decreasing in p, and use Lemma A3 to show that an increase in p leads to more concentrated attacks. ■

Empirical Appendix

In this brief empirical appendix, we introduce supplemental analyses of close combat (direct fire) and improvised explosive devices.

As we describe in the main text, insurgents have varying levels of control over the timing of their attacks. Indirect fire events can be initiated at any time against stationary targets like bases or military outposts. The timing of direct fire and IED attacks cannot be controlled unilaterally by rebels. Close combat, for example, is often characterized by attacks on convoys and non-stationary targets. Troops and vehicles are rotated from locations on non-random schedules. The timing of these attacks, therefore, is consistently less random as a consequence of government strategy, not rebel tactics. Similarly, IEDs may be emplaced hours or days before they are triggered by a passing convoy. Rank ordered, rebels have the least control over the timing of roadside bomb attacks.

We replicate the visual evidence we introduce in our main results in Figures A-1 and A-2. Notice that direct fire, over which insurgents maintain some limited control over timing, is negatively correlated with revenue but the slope is consistently flatter than for indirect fire attacks (over which rebels have unilateral timing control). For IEDs, the coefficient is effectively zero. We introduce the regression-based evidence in Tables A-1 and A-2. Although direct fire attacks are consistently negatively correlated with revenue, the magnitude of the main effect (when compared to our indirect fire results) is consistently weaker. Similarly, although our point estimates are consistently negative with respect to roadside bombs, our precision is substantially diminished.

We interpret these findings as evidence consistent with our main results. A lack of control over the timing attacks is analogous (statistically) to measurement error, which will lead to attenuation bias (towards zero).

Figure A-1: Bivariate relationship between opium revenue and p-value of randomization test of combat (direct fire attacks) in Afghanistan

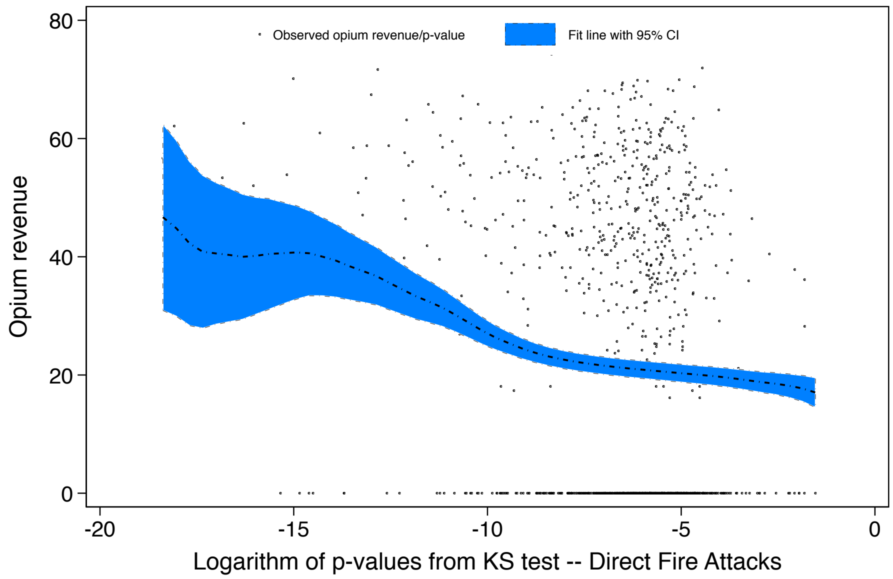


Figure A-2: Bivariate relationship between opium revenue and p-value of randomization test of combat (IED attacks) in Afghanistan

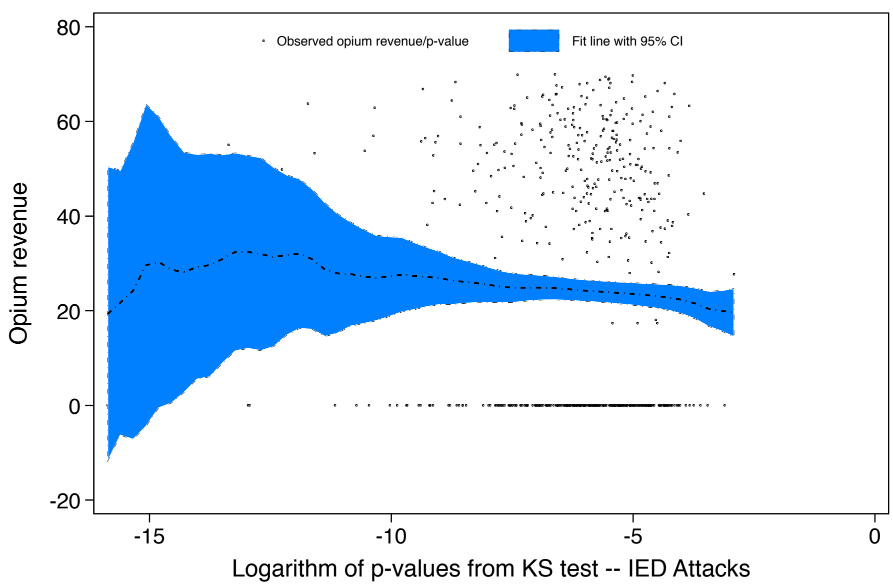


Table A-1: Impact of rebel capacity on within-day randomization of direct fire attacks

Opium Revenue	-0.0309***	-0.0110***	-0.0121***	-0.0119***
	(0.00780)	(0.00292)	(0.00328)	(0.00327)
Model Parameters				
Fighting Season FE	Yes	Yes	Yes	Yes
FS Trend		Yes	Yes	Yes
PS Trend			Yes	Yes
PLS Trend				Yes
Model Statistics				
No. of Observations	1128	1128	1128	1128
No. of Clusters	236	236	236	236
R^2	0.0963	0.448	0.450	0.464

Notes: Outcome of interest is the (log) p-value of the randomness test. The quantity of interest is opium revenue for a given district-year. All regressions include fighting season fixed effects. Column 2-4 add controls for the intensive margin of fighting during the fighting, harvest, and planting seasons respectively. Heteroskedasticity robust standard errors clustered by district are reported in parentheses. Stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A-2: Impact of rebel capacity on within-day randomization of IED attacks

Opium Revenue	-0.00849**	0.000559	-0.000232	-0.00000219
	(0.00385)	(0.00259)	(0.00280)	(0.00270)
MODEL PARAMETERS				
FIGHTING SEASON FE	Yes	Yes	Yes	Yes
FS TREND		Yes	Yes	Yes
PS TREND			Yes	Yes
PLS TREND				Yes
MODEL STATISTICS				
No. of Observations	653	653	653	653
No. of Clusters	161	161	161	161
R^2	0.0700	0.179	0.186	0.186

Notes: Outcome of interest is the (log) p-value of the randomness test. The quantity of interest is opium revenue for a given district-year. All regressions include fighting season fixed effects. Column 2-4 add controls for the intensive margin of fighting during the fighting, harvest, and planting seasons respectively. Heteroskedasticity robust standard errors clustered by district are reported in parentheses. Stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.