Online Appendix

Rebel Capacity, Intelligence Gathering, and Combat Tactics

Konstantin Sonin and Austin Wright

June 6, 2022

Appendix

Contents

Appendix A – Summary Statistics

Table A-1: Summary Statistics

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

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

Appendix B – Theory

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

Proof of Proposition 1

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

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

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

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

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

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

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

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

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

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

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

We start with a lemma.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

It is easy to check that

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

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

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

This implies that

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

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

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

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

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

where

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

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

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

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

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

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

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

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

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

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

Then, using (A6), we obtain

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

Similarly,

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

■

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

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

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

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

Let us do the following transformation:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

This concludes the proof of Proposition 1. ■

Proof of Proposition 2

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

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

Appendix C – Additional Institutional Details

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

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

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

Appendix D – Research Design

Appendix D1 – Additional measures

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

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

Appendix D2 – Covariates in the main specification and robustness checks

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

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

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

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

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

Appendix D3 – Details on temporal clustering measure

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

1. Fit a local polynomial regression to the observed distribution of violence by hour. We specify a conservative bandwidth of 1. This empirical distribution of fitted values is stored.[3]

2. Identify the sequence of district-hours during which combat engagements occur. For each district-hour, we know the sum of the number of attacks.

3. Randomly shuffle the sequence above. This is equivalent to a randomization or permutation test.

4. Fit a local polynomial regression to the randomly shuffled distribution of violence by hour. The simulated distribution of fitted values is stored.

5. Execute the bootstrap Kolmogorov-Smirnov test composed of four elements.

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

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

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

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

    (d) Calculate and store the likelihood parameter of the tests as $\sum_{b=1}^{1000} \frac{1\{T^{KS}_{dfi,b} > T^{KS}_{dfi}\}}{1{,}000}$, where the numerator is an indicator function.

6. Repeat steps 2 through 5 1000 times. Evaluate the central tendency (mean) of the likelihood parameters.

7. Replace zero values with the minimum observed non-zero rank value and calculate the log.

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

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

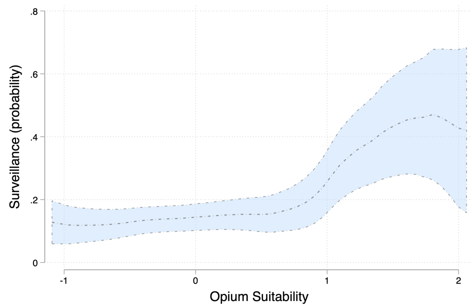
Appendix D4 – Details on spatial clustering measure

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

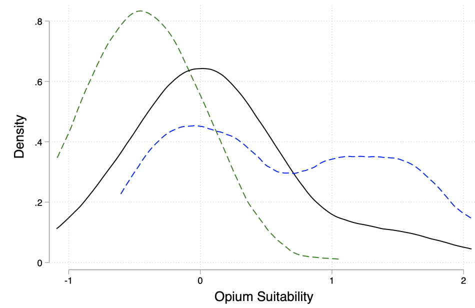
The central inferential challenge for studying spatial dynamics is identifying the set of counterfactual locations that could have been attacked but were not. If, for example, the set of locations that could be targeted was a known quantity and did not change over time, it would be possible to identify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

Appendix E – Supplemental Figures and Results

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



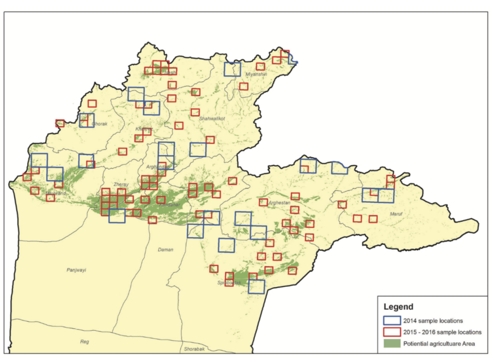
(a) Opium Suitability and Surveillance



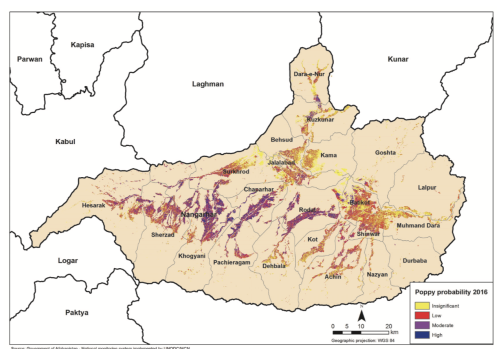
(b) Opium Suitability by Price Zone

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

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



(a) Sampling Satellite Imagery



(b) Impute Production from Imagery/Field Obs.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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