Understanding Conflict
Political Economy

Prof. Francesco Trebbi

University of British Columbia - Vancouver School of Economics

Based on:

Insurgency and Small Wars:

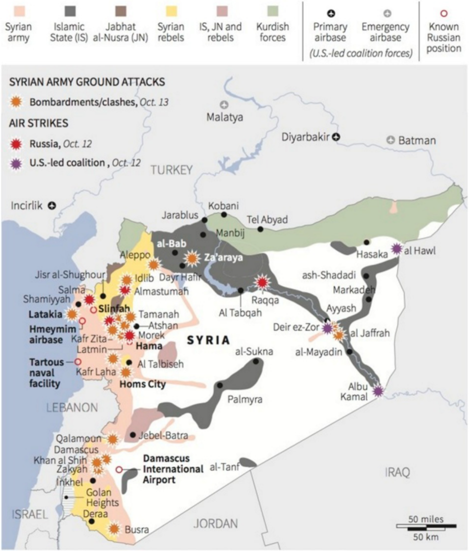
Estimation of Unobserved Coalition Structures

Francesco Trebbi - UBC, CIFAR NBER
Eric Weese - Yale University

Insurgency and Guerrilla Warfare

Insurgency & guerrilla warfare impose enormous socioeconomic costs. Often persist for decades.

- Fearon (2008): "*There are no clear frontlines*"
- Blurred lines between civilians & insurgents (Kilcullen, 2009; Berman & Matanok, 2015)
- No data. Substantial opaqueness. Severe limitations to its systematic empirical study



Airstrike locations are approximate. (Sources: Institute for the Study of War; Ministry of Defense of the Russian Federation; U.S. Central Command; Syrian Observatory for Human Rights / Reuters)

Insurgency and Guerrilla Warfare cont'd

Large body of literature in Political Science/Economics on the organization & strategies of insurgents, but still relevant questions

- **What is the organization of insurgent groups in asymmetric warfare? Are the insurgents part of a single organized group? Or part of a loosely organized umbrella coalition?**
- **Where are the insurgents?**
- **What incentives guide the insurgency (if any)?**

Do they follow/are constrained by specific economic incentives (Berman, Shapiro, Felter, 2011)? Or by ethnic constraints/cleavages (Fearon & Laitin, 2003)?

We tackle these issues in two real world insurgencies (making use of micro incident-level military data generously provided by the ESOC consortium).

Insurgency and Guerrilla Warfare: The Taliban

"The Taliban are not a unified force -they are not the SPLA in Sudan or the Maoists in Nepal"
- Ashraf Ghani, current Afghan president, lecture for the Miliband Programme at LSE, 3/12/2010 (Brahimi, 2010).

"[...] the Taliban's environment is not unified – like any other environment, that the Taliban are made of different groups. Some divide them by foreigners and nationals, some about white, gray and black; there are so many forms about it."
- Steffan De Mistura, Special Representative UN Mission in Afghanistan, Atlantic Council 7/1/2010

"The Taliban themselves are not fully united and the insurgency is not limited to the Taliban."
- Giustozzi (2009)

The Taliban (cont.)

"The Taliban are often described as an umbrella movement comprising loosely connected groups that are essentially local and unorganized. On the contrary, [...] the structure and strategy of the insurgency reveals a resilient adversary, engaged in strategic planning and coordinated action."

"The Taliban have a strategy and a coherent organization to implement it, and they have been successful so far."

- Dorronsoro (2009)

- Organizational manual for Taliban: *Obedience to the Emir*.
- Centralized code of conduct for Taliban: *Layha*

The Taliban (cont.)

"The Pakistani Taliban have different structures, different leaders, and a different social base [relative to the Afghan Taliban -AN]. *They are, in fact, an umbrella movement comprising loosely connected groups."*
- Dorronsoro (2009)

What we do

Our approach in a nutshell:

- A network of N nodes (e.g. the centroid of district) where insurgents can be present (one group or multiple ones)
- Attacks occur over time in these nodes at a specific time frequency (daily). T periods.
- $N << T$. Based on the **covariance properties of the attacks** over time it is possible to infer simultaneous control by the same group.
- We propose flexible **methodologies to estimate the number of insurgent groups, their respective geographic location, and their size**.
- We estimate structurally such parameters for the **Afghan Taliban** and for **Pakistani Insurgency**.
- This allows us to **evaluate insurgent strategy** over time (e.g. pre-2008 vs post-2008 samples in Afghanistan).

Related literature

Large literature on insurgency & small wars (see Bueno de Mesquita, 2008; Blattman & Miguel, 2010; Fearon & Laitin, 2003), but mainly focused on:

- Sparse macro data.
- Participation (opportunity cost/economic incentives).

Newer more micro literature on asymmetric conflict:

- **Micro, incident-level data** in Afghanistan & Iraq: Berman, Shapiro, Felter (2011); Berman, Callen, Felter, Shapiro (2011)
- **Civilian support for rebels:** Fair, Littman, Malhotra, Shapiro (2013); Lyall, Blair, Imai (2013); Lyall (2016)

Econometrics of networks: Bramoulle & Fortin (2009), Chandrasekhar & Lewis (2012)

Also literature on geographic concentration of industry: Ellison & Glaeser (1997, 1999).

Model Setup

- District $i \in \{1, ..., N\}$
- Time is discrete $t = 1, ..., T$ (daily frequency)
- Number of unorganized/unaffiliated violent actors in i is $\ell_i \geq 0$
- $j \in \{1, ..., J\}$ finite set of organized insurgent groups
- Number of members of group j present in location i is $\alpha_{ij} \geq 0$

Random Violence

- η is the probability that one unorganized militant launches an attack at t
- Random violence independent across unorganized militants within the districts, across districts & time.
- Expected number of daily unorganized attacks in i is $\eta\ell_i$
- Variance in the number of daily unorganized attacks in i is $\eta\left(1-\eta\right)\ell_i$
- Covariance of unorganized attacks across districts i and i' is 0

Organized Violence

- Members of an organized group are more likely to attack on some days than others.
- ε_{jt} is the probability that one j organized militant launches an attack at t
- σ^2 variance of ε
- ε_{jt} is the same for all members of j independently of district i
- See Deloughery (2013) for **use of simultaneous attacks** by terrorist groups. Examples: Southern Thailand; Tamil Tigers; Indian Mujahideen; Kurdish independentists.
  Assumption appears particularly appropriate for insurgency close to international jiadist groups.

Organized Violence

- Conditional on realized ε_{jt} probability of attack is independent of other attack decisions.
- Covariance of attacks over time among any two members of j is σ^2
- Assume $\varepsilon_{jt} \perp \varepsilon_{j't}$
- Covariance of attacks over time among a member of j & one of j' is 0

Covariance Across Districts

- Consider members of group j
- Consider two districts i, i'
- Given α_{ij} and $\alpha_{i'j}$, the covariance component of attacks over time across these two districts is $\sigma^2\alpha_{ij}\alpha_{i'j}$
- If there are multiple groups in i and i' the covariance of attacks over time across these two districts is:

$$\gamma_{ii'} =_{j=1}^{J} \sigma^2\alpha_{ij}\alpha_{i'j}$$

In model $\gamma_{ii'} \geq 0$
(in finite samples could be negative only due to sampling error)

Were α_{ij} time-varying with mobile insurgents hitting first district i & then a different district i', it would imply $\gamma_{ii'} \leq 0$

Covariance Across Districts

- Call Γ the variance-covariance matrix of attacks across districts.
- Decompose $\Gamma = \Gamma_D + \Gamma_L$
  where the matrix:

$$\Gamma_L = \sigma^2 \begin{bmatrix} \sum_j \alpha_{1j}^2 & \sum_j \alpha_{1j}\alpha_{2j} & & & \\ \sum_j \alpha_{2j}\alpha_{1j} & \sum_j \alpha_{2j}^2 & & & \\ \dots & & \sum_j \alpha_{ij}^2 & & \\ \sum_j \alpha_{ij}\alpha_{1j} & & & \dots & \sum_j \alpha_{ij}\alpha_{i'j} \\ \dots & & & & \dots \end{bmatrix} \qquad (1)$$

- Define $\overline{\Gamma}$ the sample counterpart to Γ

Trace Minimization

$$\min_{\hat{\Gamma}_D} \operatorname{Tr}(\hat{\Gamma}_L) \tag{2}$$
$$\text{s.t. } \hat{\Gamma}_L = \bar{\Gamma} - \hat{\Gamma}_D, \quad \hat{\Gamma}_D \text{ diagonal,}$$
$$\hat{\Gamma}_D \succ 0, \hat{\Gamma}_L \succ 0$$

$\hat{\Gamma}_L$ is the estimated counterpart to Γ_L
Tr() denotes the sum of diagonal entries of a matrix
$\succ 0$ indicates positive semi-definiteness

Saunderson et al. [2012] show that a sufficiently low-rank positive semi-definite matrix is recoverable

Estimation

We desire estimates for a large set of parameters:

- $\widehat{J}$ the estimated number of insurgent groups.
- Estimates of the $N \times J$ parameters $\widehat{\alpha}_{ij}$
- We employ multiple approaches: 1) **Clustering**; 2) **Non-Negative Matrix Factorization**.
- No confidence intervals available for most parameters (e.g. J is discrete), but permutation tests allow to reject some important null hypotheses.

k-means Clustering

Γ is $(N \times N)$ high dimensionality (see Luxburg 2007)

Clustering aims a producing a partition of the N districts into $J \leq N$ clusters:

- $k-$means clustering is an example: Take N objects at distance $d_{ii'}$ from each other & allocate each to one of k clusters such that each object belongs to the cluster on average closer to it (i.e. with the nearest mean from it).
- Turns out that even this simple clustering algorithm is NP-hard (all possible partitions have to be tried to find the optimum).
- Computer Scientists typically work to find the fastest algorithm/approximately optimal.
- Little attention given to asymptotic properties of clustering technique as an estimator of the underlying clustering structure.

Clustering (cont.)

We rely on a more structural clustering approach & go in more depth on the econometric properties.

Assumption

At most one insurgent group in a district.

Clustering (cont.)

Starting with at most one organized group j in i & i' $\gamma_{ii'} = \alpha_{ij}\alpha_{i'j}$

If this is true, then the covariance matrix Γ_L can be represented as the **block diagonal matrix**:

$$\Gamma_L = \sigma^2 \begin{bmatrix} \Gamma_L^1 & 0 & 0 & ... & \\ 0 & ... & & & \\ ... & & \Gamma_L^j & & \\ 0 & & & ... & 0 \\ 0 & & & & \Gamma_L^J \end{bmatrix}$$

where block j consists of only the districts containing group j. Example: if j is in districts $1, 2, 3$ only:

$$\Gamma_L^j = \sigma^2 \begin{bmatrix} \alpha_{1j}\alpha_{1j} & \alpha_{1j}\alpha_{2j} & \alpha_{1j}\alpha_{3j} \\ \alpha_{2j}\alpha_{1j} & \alpha_{2j}\alpha_{2j} & \alpha_{2j}\alpha_{3j} \\ \alpha_{3j}\alpha_{1j} & \alpha_{3j}\alpha_{2j} & \alpha_{3j}\alpha_{3j} \end{bmatrix}$$

Simple Spectral Approach

Use $\hat{\Gamma}_L$ directly: Each Γ_L^j has rank 1. Asymptotically $\hat{\Gamma}_L$ has J non-zero eigenvalues.

Thus eigenratio $\text{ER}_{j-1} = \hat{\lambda}_{j-1}/\hat{\lambda}_j \to c_j$ for $j \in \{2, ...J\}$, whereas $\hat{\lambda}_J/\hat{\lambda}_{J+1} \to \infty$ as $T \to \infty$ [Ahn & Horenstein ECMA 2013]

Using (estimated) $\hat{\Gamma}_L$:

1. Calculate $\hat{\lambda}$
2. Look at **eigenratio** $\hat{\lambda}_{k-1}/\hat{\lambda}_k$ for $k \in \{2, ...K\}$
3. Finite-sample modification: Subtract off $\tilde{\lambda}_{k-1}/\tilde{\lambda}_k$ expected from random noise. Eigenvalues of covariance matrix w/ no clusters. "No cluster" is generated by randomly rearranging the time indices of the observations for each district. Done independently for each district, the result is data with a covariance matrix solely the result of random variation.
4. Is there something that looks "large"? This is $\hat{J}$

Clustering of Districts

Block diagonal matrix Γ_L is **adjacency matrix of a weighted graph with J connected components**

1. Construct a scaled version of Γ_L:

$$\Gamma_L^{\text{cor}} = D(\sum_j \alpha_{.j}\alpha_{.j})^{-1/2}\Gamma_L D(\sum_j \alpha_{.j}\alpha_{.j})^{-1/2},$$

$$\Gamma_L^{\text{cor}} = \begin{bmatrix} 1_{N_1} & 0 & & \\ 0 & 1_{N_j} & & \\ \dots & & & \\ 0 & & \dots & 1_{N_J} \end{bmatrix}$$

   known as "sphering". It turns a covariance matrix into a correlation matrix

2. $1 - \gamma_{ii'}^{\text{cor}}$ is the cosine distance between i and i'
3. Use standard (weighted) k-means clustering algorithm with cosine distance

Estimating Group Membership

- To estimate the $\widehat{\alpha}_{ij}$ we propose an approximate estimator.
  1. Recall that we can estimate $\overline{\gamma}_{ii'} = \widehat{\alpha_{ij}\alpha_{i'j}}$ & that the sum of the rows of $\overline{\Gamma}^j$ corresponding to district i is $\sum_{i'\neq i} \widehat{\alpha_{ij}\alpha_{i'j}}$
  2. If each organized group is present in a large number of districts & no single districts has an α_{ij} too large, then it is reasonable to use the approximation

$$\begin{aligned}\sum_{i'\neq i} \alpha_{ij}\alpha_{i'j} &\simeq \sum_{i'} \alpha_{ij}\alpha_{i'j} \\ &= \alpha_{ij} \sum_{i'} \alpha_{i'j} \\ &= \alpha_{ij} a_j\end{aligned}$$

  3. Since a_j is constant across i's, the sum of the rows of $\overline{\Gamma}^j$ across i's gives the relative prevalence of group j's members across districts in block j (across all districts in case $J = 1$).

More Sophisticated Approach

More sophisticated approach to estimation of J

- **Gap Statistic** [Tibshirani, Walther, Hastie 2001]:

$$\text{Gap}(k) = E^*[W_k] - W_k$$

- Here W_k is the variation that is not explained by the k clusters = Sum of squared distances within clusters defined w.r.t. a set of covariates Z (e.g. geographic distance)
- Any systematic correlation between Z and α_{ij} will induce within **sum of squares lower under the real data than random data, if there are insurgent groups**.
- E^* is the expectation taken w.r.t. a reference distribution chosen to correspond to no cluster structure.
- "No cluster" is generated by randomly rearranging the time indices of the observations for each district. Done independently for each district, the result is data with a covariance matrix solely the result of random variation.

More Sophisticated Approach (cont.)

- Gap(k) gives the fit of the $k-$cluster structure in the actual data relative to the fit it obtains on random data.
- The estimated number of clusters $\widehat{J}$ is selected to be the smallest k such that:

$$\text{Gap}(k+1) - \text{Gap}(k) \leq s_{k+1}$$

  where s_{k+1} is the estimated standard error for the objective function, obtained by randomly drawing a large number of covariance matrices from the reference "no cluster" distribution & then calculating W_{k+1} for each of these matrices.

- **Stop when the gain in fit by adding clusters (i.e. parameters) is smaller than the likelihood of overfitting the randomly reshuffled data (i.e. the dispersion of the fit obtained by chance on "no cluster" data).**

Non-negative Matrix Factorization

Assumption

Multiple insurgent groups can be in a district.

Non-negative Matrix Factorization (cont.)

More direct approach:

- Suppose for now J is known
- Consider an estimator $\widehat{\alpha}_{ij}$ such that holds the restriction

$$\overline{\gamma}_{ii'} =_{j=1}^{J} \widehat{\alpha}_{ij}\widehat{\alpha}_{i'j}$$

- $N(N-1)/2$ such restrictions for $N \times J$ parameters
- Necessary for identification is $J \leq (N-1)/2$ (likely our case)

Non-negative Matrix Factorization (cont.)

An estimator can be expressed as

$$\underset{\hat{\alpha}_{ij}\geq 0}{\operatorname{argmin}} \sum_{i} \sum_{i' \neq i} (\overline{\gamma}_{ii'} - \sum_{j} \hat{\alpha}_{ij} \hat{\alpha}_{i'j})^2 \tag{3}$$

- Lots of parameters even for a small number of groups $J \leq 5$, but manageable. See Birgin, Martinez, Raudan (2000).

Modified Eigenratio for NMF

- An eigenratio estimator in this context is no longer valid as the rank of Γ_L is not J when groups have the potential to overlap.
- Instead of the rank of Γ_L, we thus base our estimate $\hat{J}$ on the completely positive rank of Γ_L: that is, the rank of A, where $\Gamma_L = AA^{\mathsf{T}}$, and all entries of A are non-negative.

Modified Eigenratio for NMF (cont.)

- A ratio equivalent to Ahn and Horenstein's "eigenratio" can then be expressed as

$$\mathrm{NNR}_k = \frac{||\hat{\Gamma}_L - A_k A_k^{\mathsf{T}}||_F^2 - ||\hat{\Gamma}_L - A_{k-1} A_{k-1}^{\mathsf{T}}||_F^2}{||\hat{\Gamma}_L - A_{k+1} A_{k+1}^{\mathsf{T}}||_F^2 - ||\hat{\Gamma}_L - A_k A_k^{\mathsf{T}}||_F^2} \tag{4}$$

  where $||||_F$ is the Frobenius norm.

- The intuition for NNR_k is exactly that of the eigenratio approach: if Γ_L has a completely positive rank of k, then the $k+1$th factor should not help explain Γ_L, and thus NNR_k should diverge to infinity. In contrast, values of NNR_k for $k < J$ will converge to finite values.

Robustness

1. Consider covariance of attacks only within shorter periods (a month).
2. Drop noisy districts (too few attacks)
3. For Afghanistan use only incidents explicitly coded as claimed by Taliban
4. Jackknife by dropping one month of data at a time and re-estimating the model

Data

- Worldwide Incidents Tracking System (WITS) 2004-2009
- Geocoded and available for Afghanistan & Pakistan
- Pakistan also has BFRS data:
    - twice as many incidents coded in WITS time period
    - much longer coverage

Typical **incident reported**:

- *"On 27 March 2005, in Laghman, Afghanistan, assailants fired rockets at the Governor House, killing four Afghan soldiers and causing minor damage. The Taliban claimed responsibility for the attack."*

Data

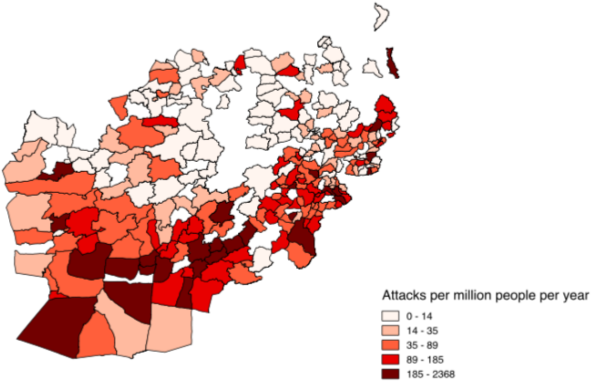
Afghanistan:

- 398 districts. 123 districts with 0 incidents 2004-2009
- Rivers, main roads, estimated population, luminosity data for all districts
- Source: Afghan Statistical Office
- Ethnic data from Atlas Nadorov Mira (obtained from GREG)

Pakistan:

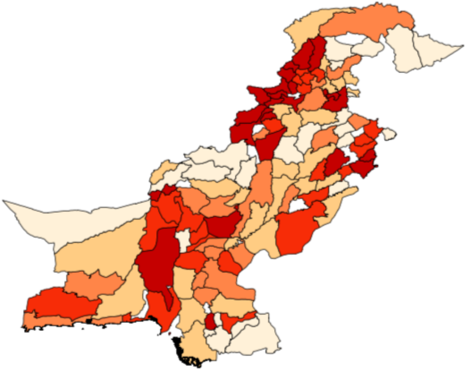
- 129/141 districts 2008-2011 (32/14 districts with 0 incidents)
- covariates in progress...

Total Incidents by District: Afghanistan

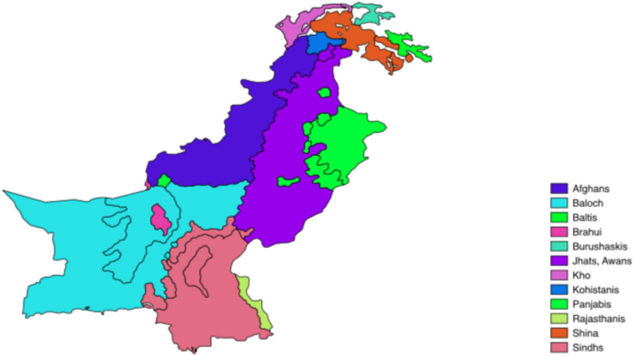


Ethnicity in Afghanistan

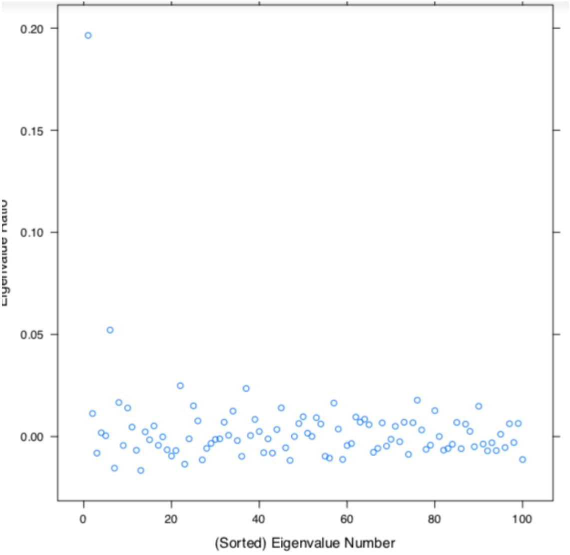
Total Incidents by District: Pakistan



Ethnicity in Pakistan



Eigenratio: Afghanistan (WITS)

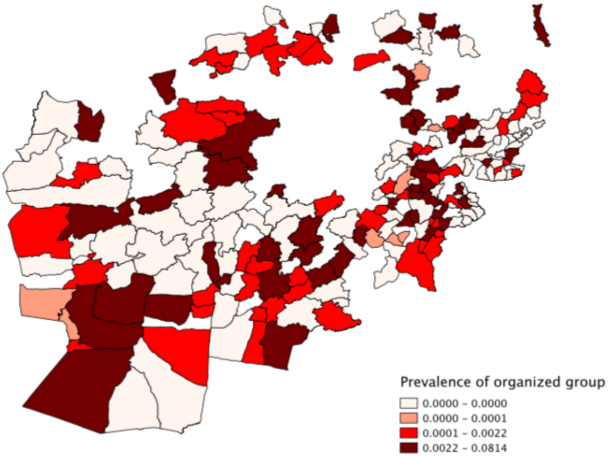


Gap Statistics Afghanistan

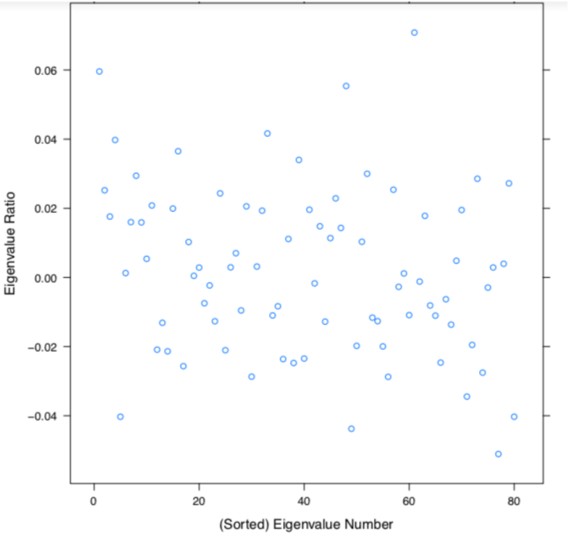
			Within-month Covariance
1 group	Randomly Reshuffled Data		16.765
	Actual Data		16.765
	Gap	A	0
2 groups	Randomly Reshuffled Data		16.760
	Actual Data		16.765
	Gap	B	-0.005
	Gap Statistic (B-A)		-0.005
	St. Dev. Randomly Reshuffled Data		0.006
3 groups	Randomly Reshuffled Data		16.755
	Actual Data		16.757
	Gap	C	-0.002
	Gap Statistic (C-B)		0.003
	St. Dev. Randomly Reshuffled Data		0.008
4 groups	Randomly Reshuffled Data		16.749
	Actual Data		16.755
	Gap	D	-0.005
	Gap Statistic (D-C)		-0.004
	St. Dev. Randomly Reshuffled Data		0.011
5 groups	Randomly Reshuffled Data		16.745
	Actual Data		16.738
	Gap	E	0.007
	Gap Statistic (E-D)		0.012
	St. Dev. Randomly Reshuffled Data		0.011

Full 2004-09 period.

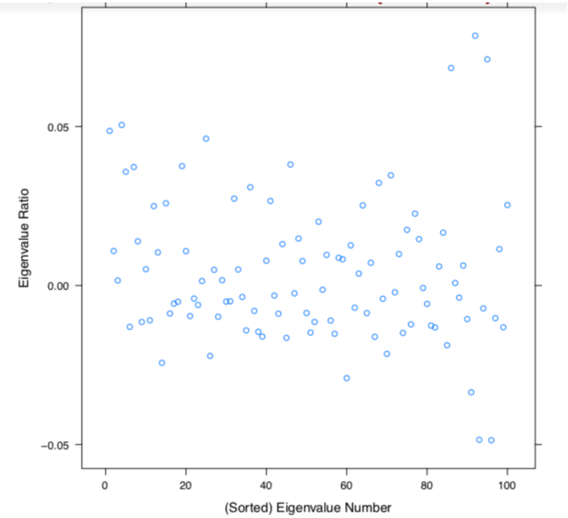
NMF Taliban Presence



Eigenratio: Pakistan (WITS)



Eigenratio: Pakistan (BFRS)

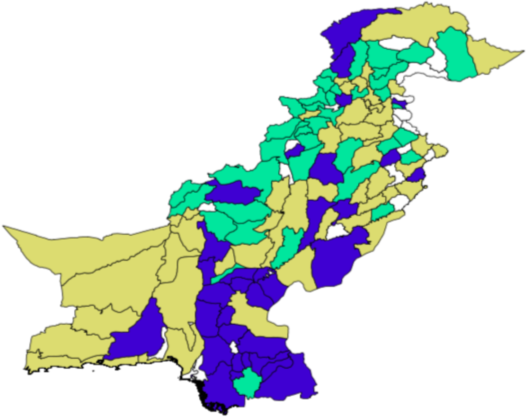


Gap Statistics Pakistan

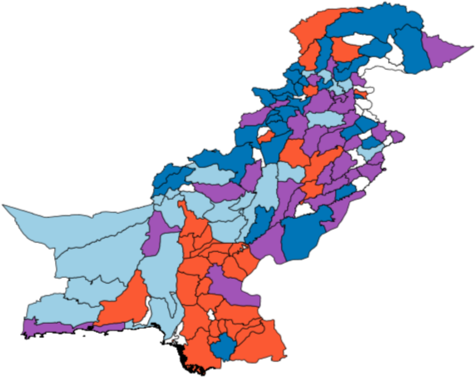
			Within-month Covariance
1 group	Randomly Reshuffled Data		16.929
	Actual Data		16.929
	Gap	A	0
2 groups	Randomly Reshuffled Data		16.921
	Actual Data		16.897
	Gap	B	0.024
	Gap Statistic (B-A)		0.024
	St. Dev. Randomly Reshuffled Data		0.011
3 groups	Randomly Reshuffled Data		16.912
	Actual Data		16.736
	Gap	C	0.176
	Gap Statistic (C-B)		0.152
	St. Dev. Randomly Reshuffled Data		0.015
4 groups	Randomly Reshuffled Data		16.903
	Actual Data		16.622
	Gap	D	0.281
	Gap Statistic (D-C)		0.105
	St. Dev. Randomly Reshuffled Data		0.019
5 groups	Randomly Reshuffled Data		16.894
	Actual Data		16.758
	Gap	E	0.136
	Gap Statistic (E-D)		-0.145
	St. Dev. Randomly Reshuffled Data		0.021

Full 2008-11 period.

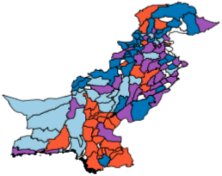
k-means: Pakistan, three groups



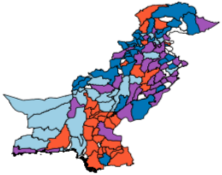
k-means: Pakistan, four groups



k-means: Pakistan, four groups: Sindhs, Panjabis



k-means: Pakistan, four groups: Baloch, Pashtun



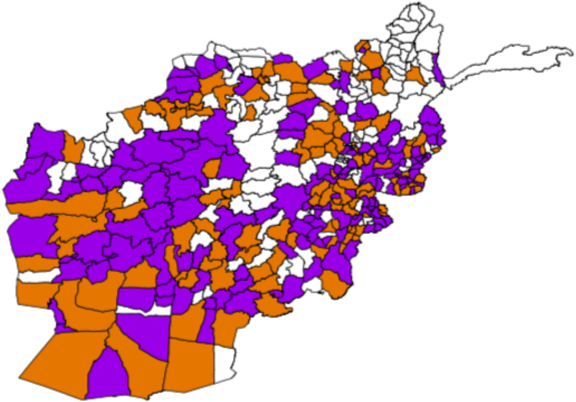
K-means and Ethnicity in Pakistan

Dep. var. is dummy for whether district has been clustered in that group

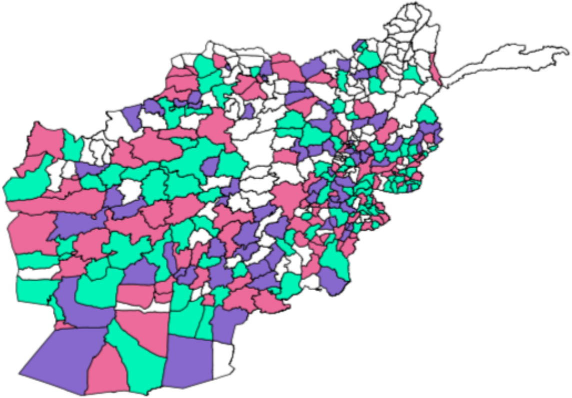
	Group 1	Group 2	Group 3	Group 4
BALOCH	0.62*** ⟵	0.25***	0.00	0.12***
	(0.09)	(0.09)	(0.10)	(0.10)
SINDHS	0.04	0.87*** ⟵	0.04	0.04
	(0.07)	(0.07)	(0.08)	(0.09)
PASHTUN	0.12*	0.08	0.58*** ⟵	0.23**
	(0.07)	(0.07)	(0.08)	(0.08)
PANJABS/JHATS/AWANS	0.16***	0.12**	0.23***	0.49** ⟵
	(0.05)	(0.05)	(0.06)	(0.06)
OTHERS	0.00	0.29**	0.57	0.14
	(0.13)	(0.13)	(0.15)	(0.16)
N	115	115	115	115

Others: Kashmiris, Gujars, Shina, Kho, Kohistanis. Robust standard errors in parentheses.

k-means: Afghanistan, two groups



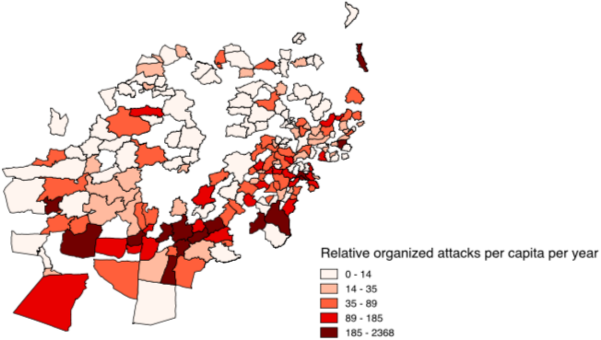
k-means: Afghanistan, three groups



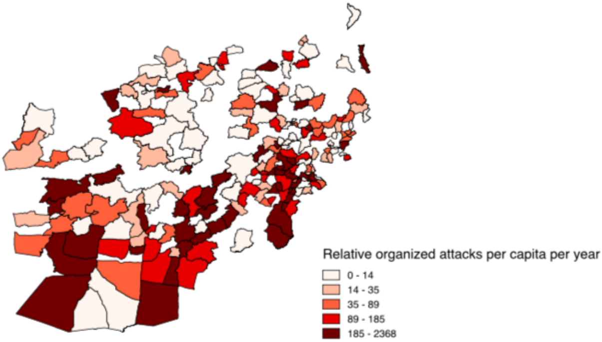
Assessing Changes in Strategy in Afghanistan

- Early data set 2004-07
- Late data set 2008-09
- Evidence of changes in insurgent control over time:
  1. Clear diffusion after 2008
  2. Penetration in non-Pashtun areas. Especially Uzbek districts.
  3. "Oil stain" strategy. See Krepinevich (2005) on Iraq.

Clustering Taliban Presence 2004-07



Clustering Taliban Presence 2008-09



NMF Taliban Presence Change

Dep. Var. is Sum of Off-Diagonal Entries of Cov. Matrix for Each District

	OLS	OLS Province FE	GLM	GLM Province FE
POST-2007	0.62	0.62	0.67	0.76
	(0.79)	(0.73)	(0.62)	(0.63)
UZBEK	-2.11	-1.05	-3.34	-2.44
	(0.30)	(0.50)	(0.67)	(0.76)
HAZARA	-1.92	-1.73	-2.18	-1.85
	(0.38)	(0.49)	(0.45)	(0.50)
TAJIK	-1.68	-0.66	-1.03	-0.57
	(0.29)	(0.47)	(0.38)	(0.48)
UZBEK*POST	1.35	1.35	2.29	2.09
	(0.52)	(0.52)	(0.80)	(0.65)
HAZARA*POST	0.19	0.19	1.50	1.55
	(0.65)	(0.53)	(0.85)	(0.70)
TAJIK*POST	0.57	0.57	0.31	0.36
	(0.46)	(0.41)	(0.50)	(0.46)
N	524	524	524	524

Conclusion

1. We **contribute methodologically to the empirical analysis of insurgency** relying on incident-level data.
2. We provide **two application to costly insurgencies**: **Afghanistan** 2004-09, **Pakistan** 2008-11
3. We find that:
   - **Afghan Taliban** present a **unified structure**
   - **Pakistan insurgents** consist of several (four) groups
   - We show how the extent and control of the **Afghan Taliban has shifted to non-Pashtun areas**
   - **Oil-spot strategy by insurgents.**

- Our approach is limited by lack of geographically finer data
  Economic costs and losses of pre-existent development project hard to quantify