

ARTICLE

Explaining Recruitment to Extremism: A Bayesian Hierarchical Case–Control Approach

Roberto Cerina¹, Christopher Barrie² , Neil Ketchley³ and Aaron Y. Zelin⁴

¹Institute for Logic, Language and Computation, University of Amsterdam, Amsterdam, The Netherlands; ²Department of Sociology, University of Edinburgh, Edinburgh, UK; ³Department of Politics and International Relations, University of Oxford, Oxford, UK; ⁴Brandeis University, Waltham, MA, USA

Corresponding author: Roberto Cerina; E-mail: r.cerina@uva.nl

(Received 26 October 2022; revised 17 February 2023; accepted 3 April 2023; published online 16 November 2023)

Abstract

Who joins extremist movements? Answering this question is beset by methodological challenges as survey techniques are infeasible and selective samples provide no counterfactual. Recruits can be assigned to contextual units, but this is vulnerable to problems of ecological inference. In this article, we elaborate a technique that combines survey and ecological approaches. The Bayesian hierarchical case–control design that we propose allows us to identify individual-level and contextual factors patterning the incidence of recruitment to extremism, while accounting for spatial autocorrelation, rare events, and contamination. We empirically validate our approach by matching a sample of Islamic State (ISIS) fighters from nine MENA countries with representative population surveys enumerated shortly before recruits joined the movement. High-status individuals in their early twenties with college education were more likely to join ISIS. There is more mixed evidence for relative deprivation. The accompanying `extremeR` package provides functionality for applied researchers to implement our approach.

Keywords: Bayesian analysis; spatial autocorrelation; rare events; multilevel modeling; extremism

Edited by: Jeff Gill

1. Introduction

Identifying who is more likely to join an extremist movement is a pressing issue for both political science and public policy. However, empirical research on this topic is beset by methodological challenges. Population surveys offer little insight into the phenomenon as recruits to extremism are tiny minorities in any society, and so are tiny minorities in samples. This is before obvious problems related to eliciting truthful responses to questions probing illicit actions. Recent innovations in survey and online digital trace methodologies have allowed researchers to obtain more accurate measures of support for extremism (Bail, Merhout, and Ding 2018; Blair *et al.* 2013; Corstange 2009; Mitts 2019). However, these approaches capture attitudes rather than behavior. For researchers interested in why some individuals join extremist movements and not others, the most common strategy is to collect a convenience sample of recruits. Using these data, scholars typically either (i) report sample proportions of a given characteristic, for example, the percentage of recruits who have college (university) education, or (ii) assign recruits to meaningful contexts and use the characteristics of those places to explain variation in the recruitment rate. While the first approach is descriptively useful, it fails to account for population baselines and other confounding factors affecting the incidence of recruitment. The second approach does provide a counterfactual and allows for multivariate analysis but suffers from familiar problems of ecological inference (Robinson 1950).

The method we propose in this paper allows researchers to leverage both survey *and* contextual data to make robust inferences about the individual and ecological correlates of recruitment to extremism. To do so, we take inspiration from the case–control design used in epidemiology and show how it can be adapted to combine a convenience sample of cases (recruits to extremism) with controls (respondents from a representative survey). In this, we build on the recent introduction of case–control methods to political science by Rosenfeld (2017; 2018), who shows how this design can be used to study protest participation and other forms of rare political behavior. Several statistical challenges arising from the nature of extremism remain, however. In particular, popular approaches for modeling rare events (King and Zeng 2001) do not account for hierarchical data structures or spatial autocorrelation in the incidence of recruitment. We also have to account for potential separation issues and the possibility of contamination between cases and controls (Rosenfeld 2018).

Our approach offers a complete solution to these statistical problems and can be described as a hierarchical, Bayesian case–control design that is robust to rare events, contamination, and spatial autocorrelation patterning the incidence of recruitment (Rota *et al.* 2013). Following Rosenfeld (2018), the Bayesian approach is preferable for a number of reasons. First, it permits the use of informative priors to account for the true prevalence of the event of recruitment, as well as to regularize coefficient estimates to account for separation bias and instability when carrying out regressions (Heinze and Schemper 2002). Second, in the absence of prior knowledge of the overall propensity of being a recruit in a given context, the model can estimate the propensity from the data (Rota *et al.* 2013). Finally, Bayesian probabilistic programming software provides unique flexibility in the modeling of the complex hierarchical structures characterizing recruitment into extremism.

A great strength of our method—and the open-source software that accompanies this paper—is that applied extremism researchers can choose those parameters most relevant for their case. When sampling from national populations, the risk of contamination between cases and controls may be sufficiently low such that it does not pose a threat to inference. On the other hand, recruitment may not qualify as a rare event when comparing recruits to certain subpopulations. So too, spatial autocorrelation in recruitment may not apply if sampling from a small area or closed context. Our modeling strategy is flexible to the inclusion or exclusion of these parameters, depending on the case at hand. In support of our approach, and to help guide the modeling decisions of future practitioners, we provide practical advice and an extensive simulation study that compares our model to alternative frameworks, and show its robustness and superiority in predicting the true underlying probability of recruitment under various bias-inducing scenarios.

To display some of the key properties of our modeling strategy, we analyze recruitment of Sunni Muslim males in nine MENA countries to the Islamic State in Iraq and Syria (ISIS). We focus our analysis on an individual's level of education and social status—two key factors associated with recruitment to extremism found in the literature on violent Islamist movements (Gambetta and Hertog 2016; Krueger 2017; Krueger and Maleckova 2003; Mesquita 2005; Morris 2020). We show how our approach can be used to perform two types of analyses. In the first, we leverage a multilevel regression model trained on a cross-national sample of ISIS recruits and non-recruits. This provides a robust descriptive analysis about the individual-level characteristics of recruits across countries and subnational administrative units. A second analysis focuses on two countries for which we have rich contextual information: Egypt and Tunisia. This analysis adds value by adjusting for local heterogeneity with the addition of relevant ecological covariates, allowing us to ascertain the potential sensitivity of individual-level findings to unobserved contextual confounding.

For the purposes of illustration, we implement the complete solution described above, accounting for spatial autocorrelation in recruitment, the possibility of contamination, and separation in our regression coefficients. Overall, we find that high-status males with college education in their early twenties were more likely to join ISIS. We also find that relatively deprived males in Egypt were more likely to join ISIS, but not in Tunisia. This heterogeneity in the individual and contextual correlates of violent extremism demonstrates the importance of accounting for both individual- and context-specific factors.

2. Explaining Recruitment to Extremism

A common strategy available to researchers interested in the correlates of recruitment to extremism is to sample on the dependent variable, obtaining relevant demographic information on individual extremists or members of extremist movements. In the ideal scenario, researchers are able to obtain movement membership lists, which can reveal information on tens of thousands of individuals (e.g., Biggs and Knauss 2012), although in practice such complete data are rare. Absent such lists, a well-established strategy is to leverage data from arrests or killings to generate samples of participants (e.g., Ketchley and Biggs 2017; Krueger and Maleckova 2003; Skare 2022). Alternatively, researchers can look to collect demographic information on extremists by either interviewing former recruits (e.g., Bérubé *et al.* 2019; della Porta 2013) or by reconstructing the biographical profiles of prominent individuals from open-source information (e.g., Gambetta and Hertog 2016; Jensen, Atwell Seate, and James 2020; Ketchley, Brooke, and Lia 2021). Per Rosenfeld (2018), a principle limitation of these samples is that they do not provide information on individuals outside of the subpopulation of interest, meaning that it is not possible to compare recruits to the population from which they are drawn. To remedy this, researchers typically either confine attention to variation among recruits (e.g., Morris 2020), or else assign individuals to meaningful contexts, for example, universities, cities, or countries, and then use the characteristics of those units to explain cross-sectional variation in the recruitment rate (e.g., Barrie and Ketchley 2018; Pape 2021). While this latter approach is undoubtedly superior to simply analyzing sample proportions, it inevitably relies on ecological inference.

2.1. A Hierarchical Bayesian Case–Control Design

In what follows, we suggest two new methods for analyzing recruitment to extremism. The first leverages a cross-national, multilevel regression model trained on a complete sample of recruits and survey respondents. This provides a robust descriptive analysis about the individual-level factors which characterize recruits across countries and subnational units. The model uses random effects to control for unobservable subnational heterogeneity; these are preferable to fixed effects due to potentially heavily imbalanced area-level sample sizes (Clark and Linzer 2015; Gelman and Hill 2006). The model further uses a conditionally autoregressive prior (Besag, York, and Mollié 1991; Morris *et al.* 2019) to account for spatial smoothing. The second analysis focuses on single country studies where rich contextual information is available. The added value of this analysis lies in controlling for local heterogeneity in order to ascertain the robustness of any individual-level findings to contextual confounding. Taken together, our proposed setup thus plots a way forward for researchers to combine survey and ecological information for the robust analysis of recruitment to extremism.

2.2. Simple Case–Control Setup

We begin by describing the backbone of our model, which is a logistic regression accounting for case–control sampling protocol via an offset. Borrowing from Rota *et al.* (2013), we define $r_i = \{0, 1\}$ as the set of states that observation i in our sample of size $n = n_0 + n_1$ can obtain, where $r_i = 1$ implies the observation is a “case”, $r_i = 0$ defines a control, $n_1 = \sum_i^n \mathbb{1}(r_i = 1)$, and $n_0 = \sum_i^n \mathbb{1}(r_i = 0)$. In our application, a “case” would refer to a known extremist; a “control” to a survey respondent. Recall that cases are selected entirely on the dependent variable, while controls come from the population that cases are drawn from. Take N_1 to represent the number of cases in the population of interest, and N_0 the number of controls. The probability of being included in the sample ($s_i = 1$) conditional on the true state of any individual can hence be understood as $P_1 = \Pr(s_i = 1 | r_i = 1) = \frac{n_1}{N_1}$, while that of being sampled as a control is $P_0 = \Pr(s_i = 1 | r_i = 0) = \frac{n_0}{N_0}$. The log ratio of these sampling probabilities can then be used as an “offset” in a logistic regression, to account for the sampling protocol. The hierarchical specification of the model follows, with regression coefficients being assigned a very weakly informative prior;¹

¹The normal distribution in our model (and in `stan`) is parameterized by mean and standard deviation. See <https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations> for prior-choice advice when using `Stan`.

$$r_i \sim \text{Bernoulli}(\rho_i), \tag{1}$$

$$\text{logit}(\rho_i) = \log\left(\frac{P_1}{P_0}\right) + \sum_k x_{i,k} \beta_k, \tag{2}$$

$$\beta_k \sim N(0, 10). \tag{3}$$

The above hierarchical model thus contains three layers: layer (1) is a model of the true state of an observation, conditional on their latent propensity ρ ; layer (2) describes this latent propensity, by accounting for systematic variation due to heterogeneity in covariates; and layer (3) models the effects of each covariate by assigning a prior probabilistic model.

2.3. Contaminated Controls

Recall that the case–control setup as described above takes known recruits and combines them with “controls” taken from survey respondents. While we know that our cases are correctly labeled, we do not know whether this is true of our controls. That is, our controls may be “contaminated” as survey respondents may have become recruits (Lancaster and Imbens 1996; Rosenfeld 2018). This is especially concerning when researchers have access to biographical information on tens of thousands of extremists (e.g., Biggs and Knauss 2012) or are comparing recruits to small subpopulations (e.g., Ketchley and Biggs 2017; Ketchley *et al.* 2021). Rota *et al.* (2013) outline a “latent variable” formulation of their contamination model. Below, we present our version of that same model as a mixture, which we find more intuitive.

The “label” of an observation, $y_i = \{0, 1\}$, is observed for all observations, while the true “state” of an observation, $r_i = \{0, 1\}$, is only observed for cases. The implied probability distribution of labels conditional on being a control is:

$$\begin{aligned} \Pr(y_i = 1 \mid r_i = 0, s_i = 1) &= 0 = \theta_0, \\ \Pr(y_i = 0 \mid r_i = 0, s_i = 1) &= 1 = (1 - \theta_0). \end{aligned}$$

Due to contamination, it is possible that observations characterized by $y_i = 0$ are actually in state $r_i = 1$; hence, we need a probability distribution for $y \mid r_i = 1$. Let $\pi = \frac{N_1}{N_1 + N_0}$ be the prevalence of recruits in the population of interest, and let $n_u = \sum_i^n \mathbb{1}(y_i = 0)$ be the number of unlabeled observations. We expect there to be πn_u cases among the unlabeled observations. We can then characterize the probability distribution of labels, conditional on being a case, as

$$\begin{aligned} \Pr(y = 1 \mid r = 1, s = 1) &= \frac{n_1}{n_1 + \pi n_u} = \theta_1, \\ \Pr(y = 0 \mid r = 1, s = 1) &= \frac{\pi n_u}{n_1 + \pi n_u} = (1 - \theta_1). \end{aligned}$$

Finally, our model for the latent state r_i must reflect the possibility of contamination. We do this by redefining the relative risk of being sampled as

$$\frac{P_1}{P_0} = \frac{\frac{n_1 + \pi n_u}{N_1}}{\frac{(1 - \pi)n_u}{N_0}} = \frac{n_1}{\pi n_u} + 1.$$

The updated, hierarchical specification for the case–control model accounting for contaminated controls is then

$$y_i \sim \text{Bernoulli}(\theta_{r_i}), \tag{4}$$

$$r_i \sim \text{Bernoulli}(\rho_i), \tag{5}$$

$$\text{logit}(\rho_i) = \log\left(\frac{n_1}{\pi n_u} + 1\right) + \sum_k x_{i,k} \beta_k, \tag{6}$$

$$\beta_k \sim N(0, 10). \tag{7}$$

In summary, we derive our labels via two distinct data-generating processes, identified by a latent state $r_i = \{1, 0\}$. In the event that the latent state of a given record is that of a true control, $r_i = 0$, it is then impossible for this record to be labeled $y_i = 1$; conversely, if the latent state is that of a true case, $r_i = 1$, then it is still possible for a record to be labeled $y_i = 0$, with probability $(1 - \theta_1)$. This latter model describes the issue of contamination. Note that in our application, θ is always observed, and fed to the model as data.

2.4. Area-Level Random Effects

Survey data and information on extremists often contain information on the origin or location of residence of individuals. We can understand individuals as nested within geographical units of increasing sizes. Generalizing, we can exploit variance at three levels: the individual, some small-area, and some large-area.

These area effects could be incorporated in the model via fixed effects, by expanding the design matrix to include relevant dummy variables for each area of interest. We consider this strategy unwise when trying to explain recruitment to extremism and prefer a random-effects approach. In the case of rare forms of political behavior, our geographical units at all levels of analysis will have relatively few observations (Gelman and Hill 2006). Additionally, for many units, we will have no cases. Finally, we know that lists of recruits are unlikely to be exhaustive; that is, we will not have data for every recruit hailing from every subnational unit or country. Here, a sample of recruitment data or similar can be treated as a non-probability sample—it is unlikely that we can have complete confidence the sample constitutes a complete or random sample of the population of interest. Given these concerns, a random-effects approach is preferable as it means: (1) we are able to borrow strength across areas, which also increases efficiency, to produce more realistic estimates for the area-level coefficients (Baio 2012; Clark and Linzer 2015) and (2) in the absence of more detailed knowledge about the data-generating process, the shrinkage effect obtained by partial pooling is more likely to shield our estimates from any systematic sampling bias among our cases (Gelman and Hill 2006).

We can also relax some of the theoretical bias associated with the shrinkage induced by random effects via incorporating observable area-level heterogeneity in the design matrix as fixed effects (Gelman and Hill 2006). This is what we elect to do in single-country analyses. Finally, it is worth highlighting that our goal is not to make inferences about area-level effects. Rather, we seek to strip our individual-level effects estimates of contested variance that may be associated with the provenance of the recruit. The resulting hierarchical model is as follows:

$$y_i \sim \text{Bernoulli}(\theta_{r_i}), \tag{8}$$

$$r_i \sim \text{Bernoulli}(\rho_i), \tag{9}$$

$$\text{logit}(\rho_i) = \log\left(\frac{n_1}{\pi n_u} + 1\right) + \sum_k x_{i,k} \beta_k + \phi_{l[i]} + \eta_{j[i]}, \tag{10}$$

$$\beta_k \sim N(0, 10), \tag{11}$$

$$\phi_l \sim N(0, \sigma_\phi), \tag{12}$$

$$\sigma_\phi = \frac{1}{\sqrt{\tau_\phi}}, \tau_\phi \sim \text{Gamma}(\epsilon, \epsilon), \tag{13}$$

$$\eta_j \sim N(0, \sigma_\eta), \tag{14}$$

$$\sigma_\eta = \frac{1}{\sqrt{\tau_\eta}}, \tau_\eta \sim \text{Gamma}(\epsilon, \epsilon), \tag{15}$$

where ϵ stands for some arbitrary number, chosen as a compromise to minimize the prior information and maximize the Markov chain Monte Carlo (MCMC) convergence speed and stability.

2.5. Spatial Autocorrelation

The network ties connecting actors across space play an important role in recruitment to high-risk activism. Sometimes the ties connecting recruits will be available; more commonly this information will not be recoverable. In the absence of detailed network information, we propose controlling for network effects at levels of varying scale. We work on the assumption that network ties are more likely to form between individuals who are geographically proximate. Depending on the richness of the data on recruits, we may generate distance matrices between geographical units of varying size.

To account for area-level spatial autocorrelation, we incorporate a version of the conditional autoregressive (CAR) model (Besag *et al.* 1991). This approach has been used in individual-level models of behavior, enabling local smoothing of predictions according to behavior observed in neighboring areas (Selb and Munzert 2011). The key ingredients of a CAR model are ω , a distance-weight matrix; α , a parameter governing the degree of autocorrelation, where $\alpha = 0$ implies spatial independence, and $\alpha = 1$ implies an intrinsic conditional autoregressive (ICAR) model (Besag and Kooperberg 1995); and σ_ψ , the standard deviation of the subnational unit effects. The resulting model for spatial random effect $\psi_l \forall l = \{1, \dots, L\}$ is then

$$\psi_l | \psi_{l'} \sim N\left(\alpha \sum_{l' \neq l} \omega_{ll'} \psi_{l'}, \sigma_\psi\right).$$

In practice, we implement the ICAR specification of the model, with $\alpha = 1$, and take ω to be the neighborhood matrix. The neighborhood matrix has diagonals zero (a unit cannot neighbor itself) and off-diagonal zero or one depending on whether the given units are neighbors. We choose this specification of the distance matrix because of the efficiency gains it affords in a Bayesian context (Morris *et al.* 2019). This leads to

$$\psi_l | \psi_{l'} \sim N\left(\frac{\sum_{l' \neq l} \psi_{l'}}{d_{l,l}}, \frac{\sigma_\psi}{\sqrt{d_{l,l}}}\right),$$

where $d_{l,l}$ is an entry of the diagonal matrix D of size $L \times L$, whose diagonal is defined as a vector of the number of neighbors of each area. The joint distribution of this model is simply a multivariate normal distribution $\phi \sim N(0, [\tau_\psi (D - W)]^{-1})$, $\tau_\psi = \frac{1}{\sigma_\psi^2}$, which is conveniently proportional to the squared pairwise difference of neighboring effects. Note that the sum-to-zero constraint is needed for identifiability, as in its absence any constant added to the ψ s would cancel out in the difference.² Following Morris *et al.* (2019), setting the precision to 1 and centering the model such that $\sum_l^L \psi_l = 0$, we arrive at

$$\log p(\psi) \propto \exp\left\{-\frac{1}{2} \sum_{l' \neq l} (\psi_l - \psi_{l'})^2\right\}.$$

The hierarchical model we implement to incorporate the spatial component is within the Besag–York–Mollié (BYM) family (Besag *et al.* 1991). For a given level of analysis, say the city or province in a cross-country analysis, BYM models are characterized by two random effects which explain unobserved heterogeneity: ϕ_l defines a non-spatial component, while ψ_l defines systematic variance due to spatial dependency. The typical challenge with BYM is that the two areal effects cannot be identified without imposing some structure since they are mutually dependent, meaning that either component is capable of accounting for contested variance at the area level. This leads to inefficient posterior exploration of any MCMC sample, and subsequent lack of convergence (Riebler *et al.* 2016). To overcome this,

²This model has the disadvantage of being an improper prior, as its density does not integrate to unity and is non-generative, though it serves our purposes within the context of a hierarchical model. The prior also encodes an intrinsic dependence between subnational units. It can no longer detect the degree of spatial autocorrelation supported by the data but instead assumes that areas are explicitly dependent, and estimates coefficients accordingly.

we implement a state-of-the-art solution leveraging penalized-complexity priors (Simpson *et al.* 2017), which proposes modeling the two effects as a scaled mixture such that

$$\gamma_l = \sigma \left(\phi_l \sqrt{(1-\lambda)} + \psi_l \sqrt{(\lambda/s)} \right),$$

where ϕ and ψ are random effects scaled to have unitary variance and $\lambda \in [0, 1]$ is a mixing parameter, defining the proportion of residual variation attributable to spatial dependency. In order for the spatial and unstructured effects to share σ , they must be on the same scale (Riebler *et al.* 2016). We must therefore scale the ICAR-distributed effects, as their original scale is defined by the local neighborhood. A proposed scaling factor is chosen such that the geometric mean of the variance parameters over the areal units is 1, $\text{Var}(\psi_l) = 1$. Note that this scaling factor, s in the equation above, can be calculated directly from the adjacency matrix, and hence it is not to be estimated but passed to the model as data.

The resulting hierarchical specification of our model follows:

$$y_i \sim \text{Bernoulli}(\theta_{r_i}), \tag{16}$$

$$r_i \sim \text{Bernoulli}(\rho_i), \tag{17}$$

$$\text{logit}(\rho_i) = \log \left(\frac{n_i}{\pi n_u} + 1 \right) + \sum_k x_{i,k} \beta_k + \gamma_{l[i]} + \eta_{j[i]}, \tag{18}$$

$$\beta_k \sim N(0, 10), \tag{19}$$

$$\gamma_l = \sigma \left(\phi_l \sqrt{(1-\lambda)} + \psi_l \sqrt{(\lambda/s)} \right), \tag{20}$$

$$\lambda \sim \text{Beta}(0.5, 0.5), \tag{21}$$

$$\phi_l \sim N(0, 1), \tag{22}$$

$$\psi_l | \psi_{l'} \sim N \left(\frac{\sum_{l' \neq l} \psi_{l'}}{d_{l,l}}, \frac{1}{\sqrt{d_{l,l}}} \right), \tag{23}$$

$$\sigma \sim \frac{1}{2} N(0, 1), \tag{24}$$

$$\eta_j \sim N(0, \sigma_\eta), \tag{25}$$

$$\sigma_\eta = \frac{1}{\sqrt{\tau_\eta}}, \tau_\eta \sim \text{Gamma}(\epsilon, \epsilon), \tag{26}$$

where $\frac{1}{2}N$ denotes a half-normal distribution, which is the recommended prior for the variance of BYM effects (Morris *et al.* 2019).

2.6. Regularizing Prior Coefficients

Multiple contributions have highlighted problems with logistic regression coefficient estimates under rare events (King and Zeng 2001). The intuition behind these challenges is typically described as some variation on the standard separation problem where any given covariate or simple combination thereof perfectly separates cases from controls. This leads to biased and unstable point estimates with large associated uncertainty (Heinze 2017). A number of regularization techniques have been proposed to reduce bias and stabilize the coefficient estimates. Our preferred regularization method is that proposed by Gelman *et al.* (2008) and Ghosh, Li, and Mitra (2018). The approach assumes that it should be unlikely to observe unit changes in the (standardized) covariates that would lead to outcome changes as large as 5 points on the logit scale. Using a slight variation on this approach to ensure sufficient regularization, we use a Cauchy prior with scale-parameter set to 1 for the regression coefficients, and a “looser” scale of 10 logit points on the intercept to accommodate for the rarity of the event in the sample. The advantages of the Cauchy prior lie in its fat tails, which avoid over-shrinkage of large coefficients (Ghosh *et al.* 2018).

We apply this prior to our fixed effects exclusively, as the likelihood of our random effects is already structured and penalized. Our final model specification is then as follows:

$$y_i \sim \text{Bernoulli}(\theta_{r_i}), \quad (27)$$

$$r_i \sim \text{Bernoulli}(\rho_i), \quad (28)$$

$$\text{logit}(\rho_i) = \log\left(\frac{n_1}{\pi n_u} + 1\right) + \sum_k x_{i,k} \beta_k + \gamma_{l[i]} + \eta_{j[i]}, \quad (29)$$

$$\beta_1 \sim \text{Cauchy}(0, 10), \quad (30)$$

$$\beta_k \mid k > 1 \sim \text{Cauchy}(0, 1), \quad (31)$$

$$\gamma_l = \sigma \left(\phi_l \sqrt{(1-\lambda)} + \psi_l \sqrt{(\lambda/\epsilon)} \right), \quad (32)$$

$$\lambda \sim \text{Beta}(0.5, 0.5), \quad (33)$$

$$\phi_l \sim N(0, 1), \quad (34)$$

$$\psi_l \mid \psi_{l'} \sim N\left(\frac{\sum_{l' \neq l} \psi_{l'}}{d_{l,l}}, \frac{1}{\sqrt{d_{l,l}}}\right), \quad (35)$$

$$\sigma \sim \frac{1}{2} N(0, 1), \quad (36)$$

$$\eta_j \sim N(0, \sigma_\eta), \quad (37)$$

$$\sigma_\eta = \frac{1}{\sqrt{\tau_\eta}}, \quad \tau_\eta \sim \text{Gamma}(\epsilon, \epsilon). \quad (38)$$

2.7. Simulation and Practical Advice

In Section D of the Supplementary Material, we outline an extensive simulation study demonstrating the performance advantage of a hierarchical Bayesian case-control approach relative to competing strategies such as the King and Zeng model (2001), as well as a simple fixed-effects logistic regression. In the simulation study, we explicitly test the performance of our model under varying values for the following parameters: (a) sample size (n); (b) population prevalence (π); (c) discrepancy between sample and population prevalence ($\pi - \hat{\pi}$); and (d) spatial autocorrelation (as measured by Moran's I). Two dimensions of our modeling framework remain untested: (i) the sensitivity of the model to poor prior information about π , the population prevalence assumed for the contamination layer, and (ii) the model's ability to deal with non-probability samples resulting from exogenous selection effects (i.e., beyond the "selection on the dependent variable" type). In Section E of the Supplementary Material, we provide actionable advice for researchers and discuss how these untested dimensions may affect the robustness of the model, in light of the results from the simulation study and the robust modeling framework we have adopted.

3. Who Was More Likely to Join ISIS?

To illustrate our approach, we analyze a set of leaked border documents capturing recruitment to ISIS. This leak was widely covered in international news media and has been used to provide descriptive statistics on the geographical distribution and demographic characteristics of ISIS fighters from multiple MENA countries (Devarajan *et al.* 2016; Sterman and Rosenblatt 2018; Zelin 2018). For the case-control design, we combine individual-level ISIS recruitment data with a nationally representative sample of Muslim males from Wave III of the Arab Barometer (2014) survey. The fieldwork for the Arab Barometer surveys was completed *before* most recruits recorded in our border documents entered ISIS-held territory, and so may be vulnerable to contamination.³

³See the Supplementary Material for more information on these data.

Our choice of covariates to use from this survey is constrained by the information included in the border documents. We elect to include covariates for age, age squared, marital status, college education, and student status. We also combine two variables for unemployed and employment in agricultural or manual labor to create a composite variable designed to measure “low status” activity. An interaction between this variable and our college education variable is designed to capture relative deprivation, that is, whether highly educated individuals engaged in low status economic activity are more likely to become recruits. Full details of each covariate are listed in the Supplementary Material.

A first model—which we refer to as the “Bird’s Eye” approach—uses a multilevel regression model trained on the complete sample of 1,051 recruits and 5,093 unlabeled records. This first model provides a robust descriptive analysis of the individual-level factors characterizing recruits across countries and subnational units.

A second model—which we refer to as the “Worm’s Eye” approach—incorporates contextual information for Egypt ($n_1 = 66, n_0 = 551$ complete records) and Tunisia ($n_1 = 426, n_0 = 589$ complete records) at the district level. We focus on these two countries due to the availability of contextual information at the district level that is not accessible for the other countries in our sample. The added value of this analysis lies in controlling for observable district-level heterogeneity in order to ascertain the robustness of any individual-level findings to contextual confounding. For both Egypt and Tunisia, we include variables to capture subnational differences in demographic and labor-market composition, employment opportunities, as well as more context-specific variables designed to capture support for Islamist political organizations and prehistories of contentious politics. Full details of all covariates are listed in the Supplementary Material.

For the main analyses, we present (1) the posterior density of fixed and random effects according to our models and (2) the posterior predictive distribution across potential recruitment profiles.⁴

3.1. Fixed and Random Effects

Figure 1 presents the posterior density of the individual-level fixed effects in the Bird’s Eye model; Figure 2a and b presents the Worm’s Eye equivalent. These plots contain the main results of our models. Note that all the covariates, including dummies, are centered and scaled; hence, the coefficients are to be interpreted in terms of standard deviations from the mean of each covariate (Supplementary Figures G.1–G.3 are the individual-level posterior densities on the original, non-standardized scale). Since we are principally interested in the robust estimation of individual-level predictors, we display only the posterior density of individual fixed effects for all of our models.⁵

The estimated intercepts for the three models are extremely low. For the Bird’s Eye model, the log odds are in the order of -11 . For the Egypt Worm’s Eye model, it is just over -13 ; in Tunisia, it is -9 . The size of the intercept is primarily driven by the size of the offset, which is in turn determined by the overall prevalence of recruitment. It is therefore not surprising that Egypt’s intercept is so dramatically low, given the close-to-zero prevalence of recruitment when compared to population size ($\pi = \frac{4}{100,000}$) versus Tunisia where this prevalence is higher ($\pi = \frac{2}{1,000}$). For the Bird’s Eye model, a different offset is provided for observations coming from different countries, to account for country-specific prevalence. The large and negative intercept underscores an important challenge in the explanation of why individuals join movements like ISIS: a linear combination of features capable of pushing an individual to become a recruit has to be extremely large, on the log-odds scale, to meaningfully affect the otherwise extremely low probability of recruitment.

⁴Convergence diagnostics are in the Supplementary Material.

⁵Supplementary Figure G.4a and b displays the standardized district-level posterior densities, whereas Supplementary Figure G.5a and b presents district-level coefficients on the original, non-standardized scale.

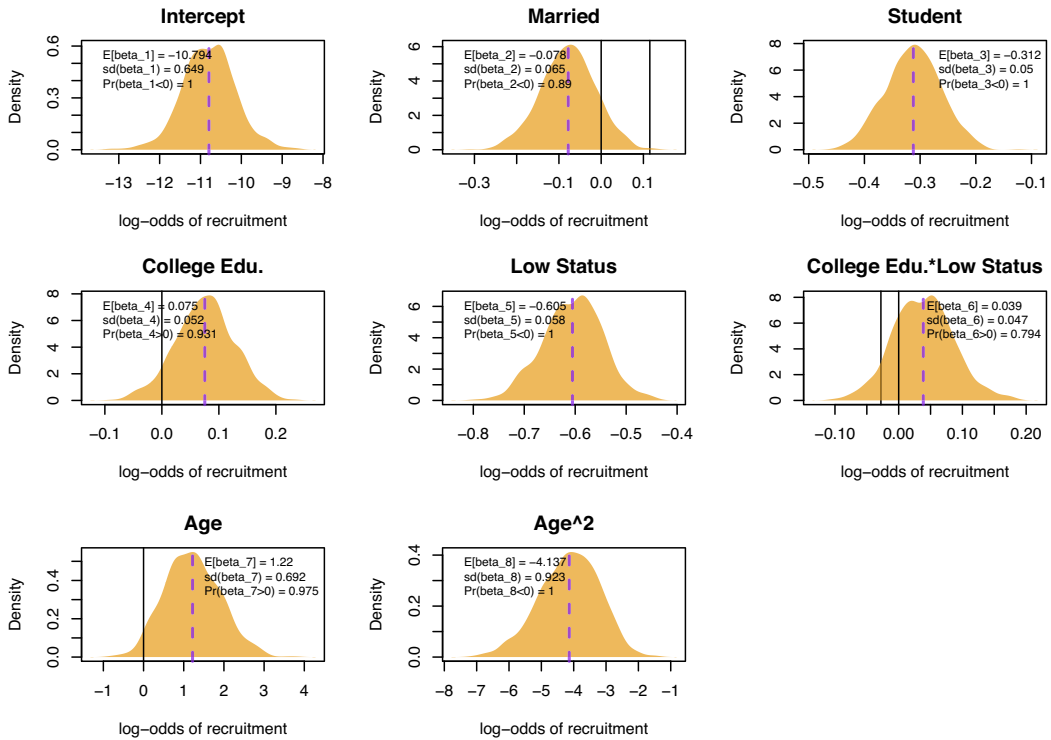
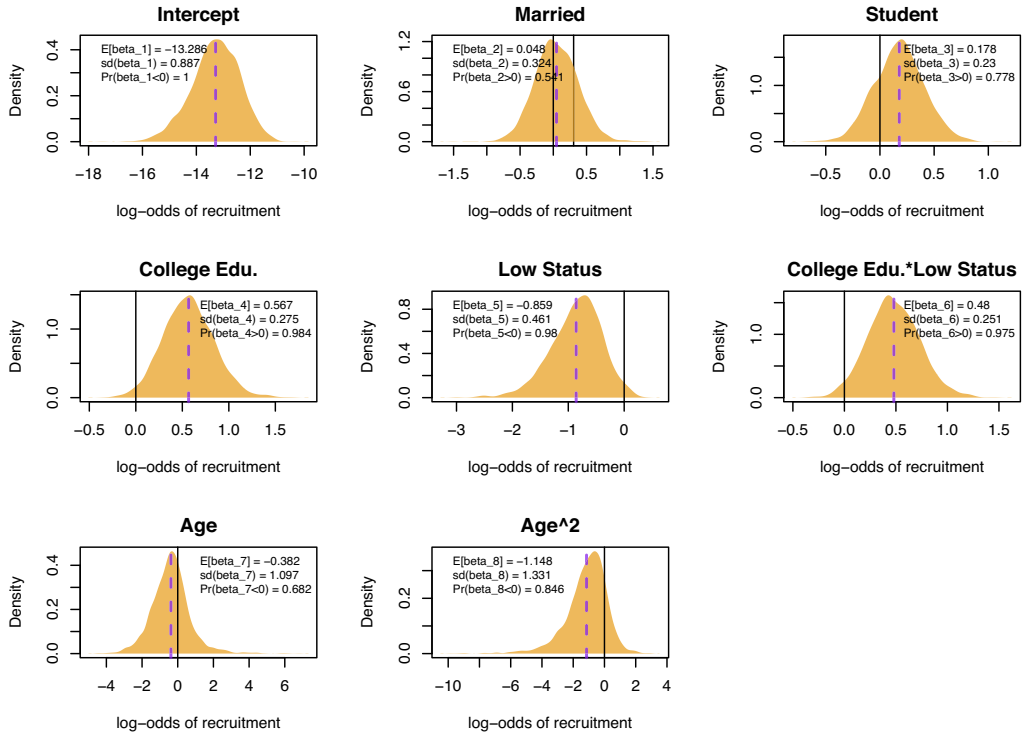


Figure 1. Posterior density of fixed-effect coefficients for the *Bird's Eye* model.

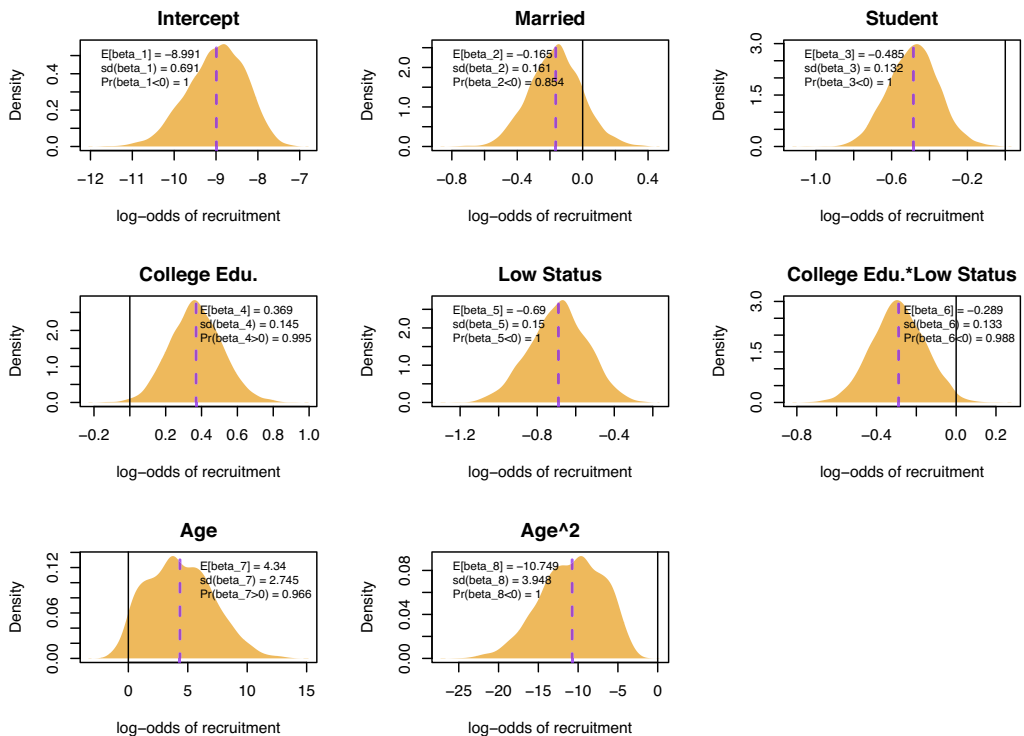
We focus primarily on testing the role of education and social status in an individual's decision to join ISIS. An individual who has college education and low status is assumed to be relatively deprived. We compare predicted log odds, as opposed to predicted probabilities, as these are scarcely comparable due to the powerful effect of the intercept, which drags probabilities of most profiles close to zero (though see Supplementary Figures H.1–H.3 for predicted probabilities of recruitment relative to the “average” profile, and Supplementary Figures H.4–H.6 for expected counts under different relative-deprivation scenarios). The total logit effects on probability of recruitment for different relative-deprivation profiles are shown in Figure 3 for the *Bird's Eye* model, and in Figure 4a and b for the *Worm's Eye*.

Relative deprivation finds mixed support: at the *Bird's eye* level, we find being high status plays a key role in increasing propensity of being recruited, while having college education plays a more minor role. A similar pattern is evident in Tunisia, though the effect of being high status and having college education is starker, meaningfully increasing the propensity to join ISIS by around 3 points on the log-odds scale compared to relatively deprived individuals. In Egypt, the effects are more consistent with relative deprivation; however, note the large prediction intervals around the total effects of relatively deprived individuals. There is also substantial overlap between the distributions in all plots. This is largely due to the uncertainty around the intercept, which plays a role in marginalizing these effects. Note further that varying prediction intervals on the effects reflect the highly unbalanced prevalence of the groups in our study. All in all, the evidence from these analyses suggests that high-status individuals were more likely to be recruited by ISIS, and that being high status *and* having a college education further increases the likelihood of recruitment. The large prediction intervals, which result from uncertainty around the intercept, underscore that much remains unknown about the underlying systematic determinants of recruitment.

To fit the ICAR model, we implemented the fully connected graph shown in Figure 5a. The spatially autocorrelated component dominates the governorate-level variance, as shown by the posterior of



(a) Egypt



(b) Tunisia

Figure 2. Posterior density of fixed-effect coefficients for the *Worm's Eye* models.

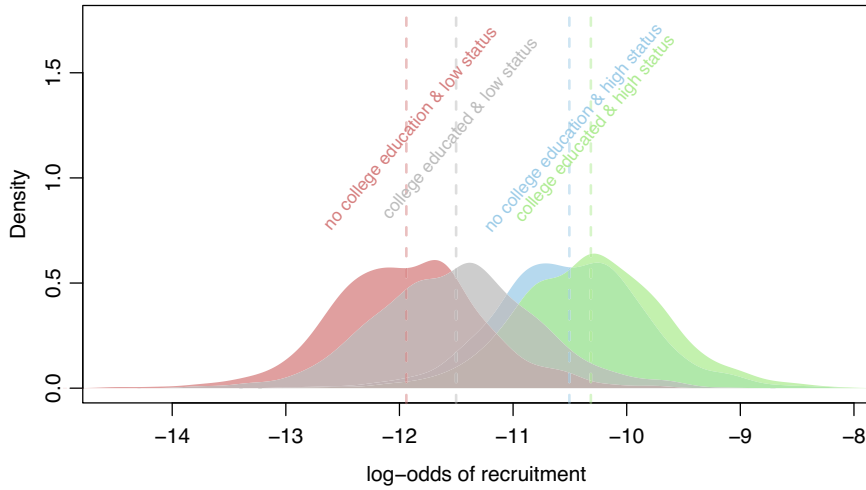
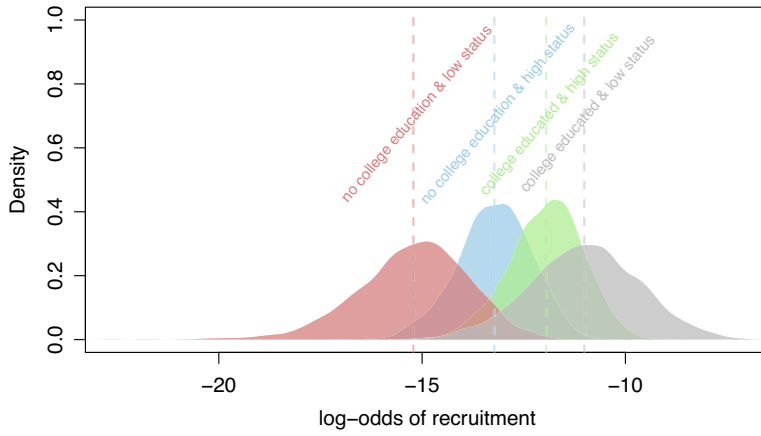
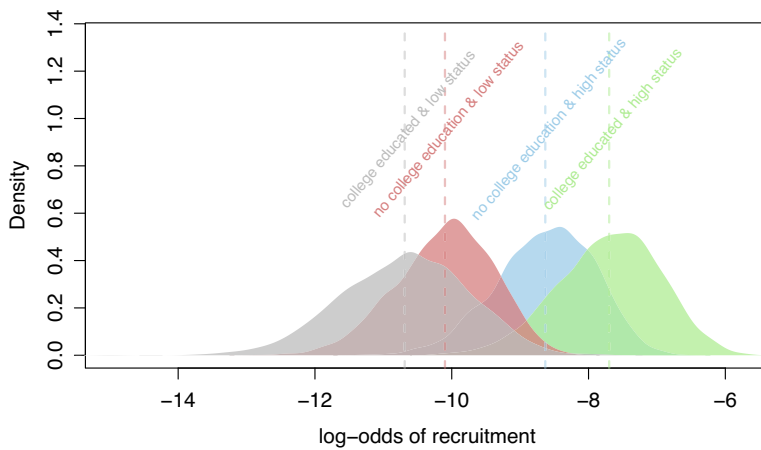


Figure 3. Predicted propensity of recruitment for relative-deprivation profiles according to the *Bird's Eye* model.



(a) Egypt



(b) Tunisia

Figure 4. Predicted propensity of recruitment for relative-deprivation profiles according to the *Worm's Eye* models.

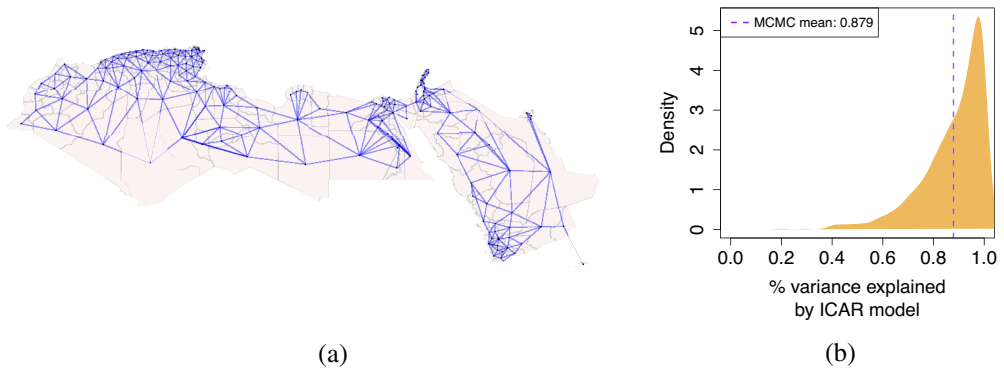


Figure 5. Fully connected graph for the Bird's Eye model (a) and Governorate-level variance mixing parameter— λ (b).

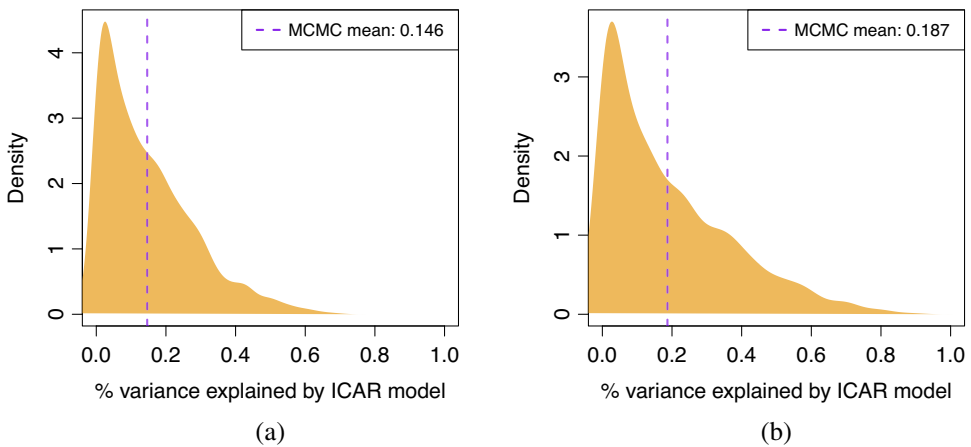


Figure 6. District-level variance mixing parameter— λ —for Egypt (a) and Tunisia (b).

mixing parameter λ in Figure 5b, estimated via Monte Carlo mean at close to 0.9, suggesting that around 90% of the variance at the governorate level can be explained by the ICAR model.⁶

We repeat these analyses for Egypt and Tunisia. Figure 6 shows similar mixing among spatial and non-spatial components for the two countries, with around 15% of the district-level variance in Egypt being explained by spatial patterns, and 19% in Tunisia. It is noteworthy that very few of our contextual variables have explanatory power for predicting recruitment. Coupled with the low percentage of variance being explained by the spatial components, our analysis suggests that, in spite of our best efforts to account for observable heterogeneity, there exist a vast array of unobserved, non-spatial district-level effects, which accounts for over 80% of the unexplained district-level variance in both Egypt and Tunisia. Hence, this contextual variance, while properly accounted for, remains unexplained. In the Supplementary Material, we also describe Moran's *I* statistics for the Worm's Eye analysis as well as point estimates for the district and governorate effects in Egypt and Tunisia (Supplementary Figure I.1).

3.2. Predicted Propensity of Recruitment by Profile

To conclude our analysis, we present inferences derived from the posterior predictive distribution of the out-of-sample probability of recruitment, focusing on individual-level characteristics.

⁶The spatial distribution of point estimates for governorate and country-level random effects are presented in Supplementary Figure I.3.

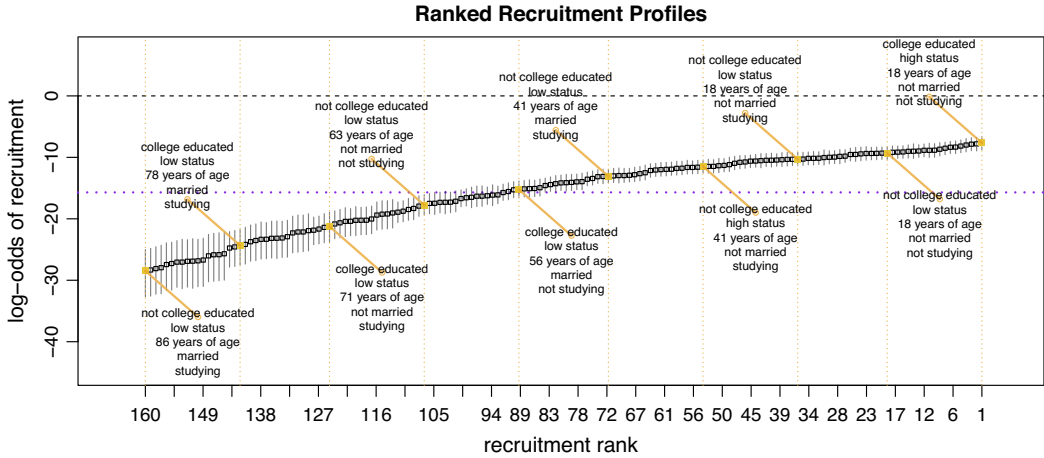


Figure 7. Distribution of the predicted probabilities prediction intervals, presented on the log-odds scale to aid cross-profile comparisons. The black dashed line highlights the zero-log-odds point, whereas the purple dotted line notes a central estimate for the median recruitment propensity across profiles.

Table 1. Top 10 recruitable theoretical profiles according to the Bird’s eye model. Profiles are ordered by predicted probability of recruitment net of sampling protocol. Ten ages are evaluated, starting at 18 (to avoid non-existent profiles) and ending at the largest observed age (86). The last four columns represent, respectively, (i) the predicted probability of recruitment, (ii) the predicted rate of recruitment per 10,000 people, (iii) the predicted odds of recruitment, relative to the “average” profile, and (iv) the log odds of recruitment.

Rank	Married	Student	College edu.	Low status	Age	$\hat{P}(r = 1 X = x)$	Predicted rate	$\frac{\hat{P}(r=1 X=x)}{\hat{P}(r=1 X=\bar{x})}$	$\text{logit}(\hat{P}(r = 1 X = x))$
1	0	0	1	0	18	0.000460	5/10,000	22.561	-7.684
2	1	0	1	0	18	0.000398	4/10,000	19.273	-7.828
3	0	0	0	0	18	0.000376	4/10,000	18.632	-7.885
4	1	0	0	0	18	0.000324	3/10,000	16.013	-8.036
5	0	0	1	0	26	0.000281	3/10,000	13.745	-8.177
6	1	0	1	0	26	0.000241	2/10,000	11.763	-8.331
7	0	0	0	0	26	0.000230	2/10,000	11.287	-8.379
8	1	0	0	0	26	0.000197	2/10,000	9.695	-8.530
9	0	1	1	0	18	0.000168	2/10,000	8.295	-8.690
10	1	1	1	0	18	0.000147	1/10,000	7.211	-8.828

What is the profile of individuals “at risk” of recruitment to ISIS according to our models? We attempt to answer this question by analyzing the predicted probabilities of all possible theoretical profiles, defined by the individual-level characteristics available in our data. Every profile is assumed to come from a hypothetical “average district”. Figure 7 presents point estimates and prediction intervals for the log odds of recruitment, over 160 possible profiles in the Bird’s eye model. Similar plots displaying the absolute and relative probabilities of recruitment are available in Supplementary Figures I.14 and I.15. Table 1 presents the top 10 most likely profiles to be recruited, providing four useful metrics to interpret the results: predicted probability; predicted rate per 10,000 people; predicted odds relative to the average profile; and log odds.

A note of caution on the interpretation of these visuals: these are useful summaries of the data, but the uncertainty around the point estimates tends to be relatively large. Taking Figure 7 as an example, a

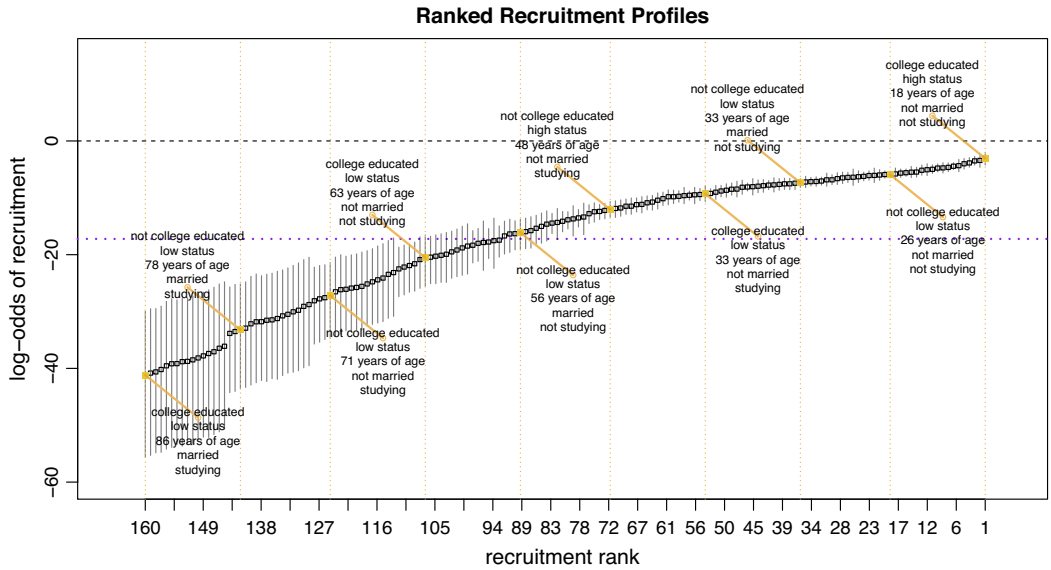


Figure 8. Worm’s Eye (Tunisia) distribution of the predicted probabilities prediction intervals, presented on the log-odds scale to aid cross-profile comparisons. The black dashed line highlights the zero-log-odds point, whereas the purple dotted line notes a central estimate for the median recruitment propensity across profiles.

qualitative interpretation of the uncertainty would be as follows: *it cannot be categorically ruled out that the most likely profile is actually ranked only 30th (out of 160), though this would be very unlikely given the evidence implied by the data.* In general, we note that profiles which are at high risk of recruitment are endowed with higher levels of certainty around their point estimates, suggesting that: (i) it is possible to distinguish high-risk profiles from low-risk profiles (at least in Tunisia and in the Bird’s eye view) and (ii) it is easier to distinguish between different high-risk profiles than it is between low-risk profiles. For Egypt, although we do observe a reduction in uncertainty at high levels of risk, we cannot entirely distinguish between low-risk and high-risk profiles, as a significant degree of overlap between posterior distributions is maintained across profiles. This is likely as a result of the relatively small sample of cases, and the large effect of the unexplained intercept.

From the Bird’s Eye prediction intervals, we notice that the predicted probability of recruitment is centered around -15 on the log-odds scale, again underscoring the rarity of becoming a recruit. A select number of profiles approach a predicted probability around -7 , and translate to meaningful rates of recruitment; these are highlighted in the predicted probabilities table, which show the 10 most recruitable profiles. Looking at Table 1, we can say that the most likely recruit profile (loosely characterized as a young, high-status, Sunni male with some college education who is unmarried and not currently studying) is around 23 times as likely to be recruited as an average Sunni male from an average area in the MENA. For every 10,000 members of the most recruitable profile across the region, we expect five to have joined ISIS. It is worthwhile to note that, consistent with Figure 3, all the most recruitable profiles are high-status individuals, and a majority of them has some college education. Unsurprisingly, all of these profiles are under 25, and not currently studying.

The Bird’s Eye profiles are comparable to the Worm’s Eye profiles for Tunisia (Figure 8 and Table 2), whereas the Egypt analysis points to stronger evidence for the relative deprivation hypothesis. In Egypt, a majority of the likely recruit profiles are relatively deprived (Figure 9 and Table 3).⁷ The relative recruitment likelihood of the most susceptible profiles in Egypt and Tunisia is also greater. In Egypt, the most likely recruit profile (loosely characterized as a young, low-status, Sunni male with some college

⁷For absolute and relative probabilities of recruitment from the Worm’s eye models, see Supplementary Figures I.16–I.19.

Table 2. Top 10 recruitable theoretical profiles according to the Tunisia “Worm’s Eye” model. Profiles are ordered by predicted probability of recruitment net of sampling protocol. Ten ages are evaluated, starting at 18 (to avoid non-existent profiles) and ending at the largest observed age (86). The last four columns represent, respectively, (i) the predicted probability of recruitment, (ii) the predicted rate of recruitment per 10,000 people, (iii) the predicted odds of recruitment, relative to the “average” profile, and (iv) the log odds of recruitment.

Rank	Married	Student	College edu.	Low status	Age	$\hat{P}(r = 1 X = x)$	Predicted rate	$\frac{\hat{P}(r=1 X=x)}{\hat{P}(r=1 X=\bar{X})}$	$\text{logit}(\hat{P}(r = 1 X = x))$
1	0	0	1	0	18	0.043680	437/10,000	335.663	-3.086
2	1	0	1	0	18	0.031493	315/10,000	242.502	-3.426
3	0	0	1	0	26	0.028925	289/10,000	222.013	-3.514
4	1	0	1	0	26	0.021177	212/10,000	159.752	-3.833
5	0	0	0	0	18	0.017397	174/10,000	137.069	-4.034
6	1	0	0	0	18	0.012615	126/10,000	98.539	-4.360
7	0	0	0	0	26	0.011647	116/10,000	88.370	-4.441
8	0	0	1	0	33	0.009503	95/10,000	73.103	-4.647
9	0	1	1	0	18	0.008662	87/10,000	68.471	-4.740
10	1	0	0	0	26	0.008327	83/10,000	64.331	-4.780

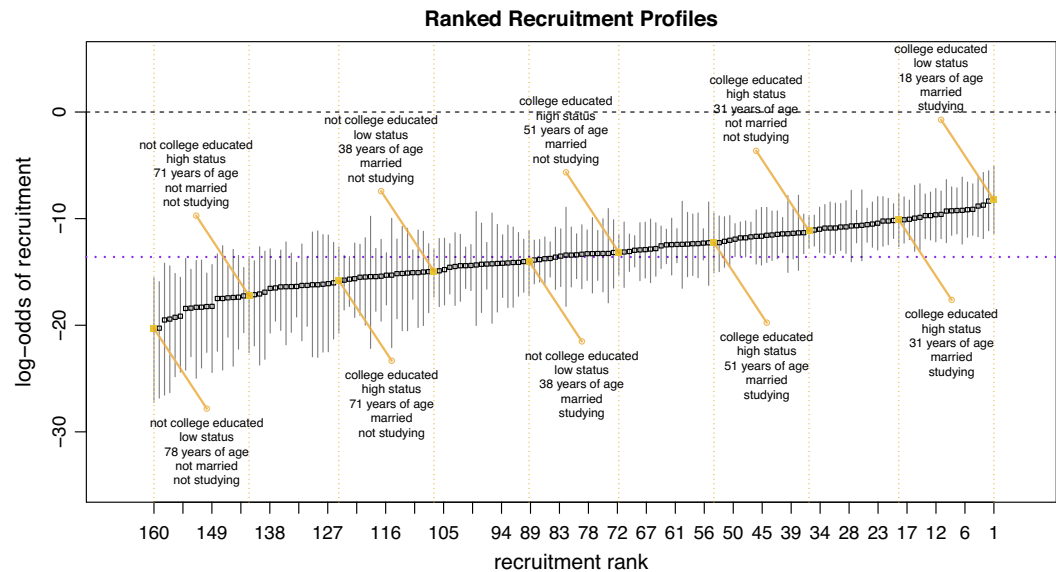


Figure 9. Worm’s Eye (Egypt) distribution of the predicted probabilities prediction intervals, presented on the log-odds scale to aid cross-profile comparisons. The black dashed line highlights the zero-log-odds point, whereas the purple dotted line notes a central estimate for the median recruitment propensity across profiles.

education who is married and is currently studying) is around 157 times as likely to be recruited as the average Egyptian Sunni male. The Egypt-specific recruitment propensity is dramatically lower than that of Tunisia, again highlighting the role of contextual effects. In Tunisia, the most likely recruit profile (loosely characterized as a young, high-status, Sunni male who has college education is unmarried and is not currently studying) has a probability of recruitment equivalent to 0.04. This profile is over 335 times as likely as the average Tunisian Sunni male to be recruited, highlighting that though recruitment is still relatively rare in the population, the probability of recruitment is far greater in the top recruitment profiles. Figure I.19 shows that only a handful of profiles have predicted probabilities above $\frac{1}{100}$.

Table 3. Top 10 recruitable theoretical profiles according to the Egypt “Worm’s Eye” model. Profiles are ordered by predicted probability of recruitment net of sampling protocol. Ten ages are evaluated, starting at 18 (to avoid non-existent profiles) and ending at the largest observed age (78). The last four columns represent, respectively, (i) the predicted probability of recruitment, (ii) the predicted rate of recruitment per 10,000 people, (iii) the predicted odds of recruitment, relative to the “average” profile, and (iv) the log odds of recruitment.

Rank	Married	Student	College edu.	Low status	Age	$\hat{P}(r = 1 X = x)$	Predicted rate	$\frac{\hat{P}(r=1 X=x)}{\hat{P}(r=1 X=x)}$	$\text{logit}(\hat{P}(r = 1 X = x))$
1	1	1	1	1	18	0.000287	3/10,000	156.721	-8.155
2	0	1	1	1	18	0.000258	3/10,000	141.021	-8.264
3	1	1	1	1	25	0.000173	2/10,000	97.486	-8.660
4	0	1	1	1	25	0.000155	2/10,000	87.536	-8.772
5	1	0	1	1	18	0.000113	1/10,000	62.422	-9.091
6	1	1	1	0	18	0.000112	1/10,000	62.529	-9.096
7	1	1	1	1	31	0.000108	1/10,000	59.223	-9.130
8	0	0	1	1	18	0.000101	1/10,000	56.978	-9.199
9	0	1	1	0	18	0.000098	1/10,000	53.336	-9.229
10	0	1	1	1	31	0.000097	1/10,000	52.977	-9.244

4. Conclusion

Extreme forms of political behavior are rarely ever committed by more than a tiny subsection of any given national population. Despite their small size, these groups often have an outsized influence on state and international politics. *Because* of their small size, extremists are particularly hard to study using conventional statistical methods and research designs.

To address this, we propose that extremism researchers take inspiration from epidemiology and recent applications of case-control methods in political science (Rosenfeld 2018). Here, we propose a new variant of the case-control design that allows us to combine survey techniques with ecological forms of analysis, allowing for meaningful comparisons with the underlying populations from which recruits are drawn. To implement this, we solve a number of statistical problems when explaining rare and extreme forms of political behavior. In particular, we demonstrate (1) how best to incorporate area-level random effects when the number of recruits for a given unit is small, (2) how to account for spatial autocorrelation in this setup, and (3) how to regularize coefficients to guard against separation. Simulations demonstrate the performance advantage of this new approach over alternatives.

While our analysis focuses on recruitment to ISIS, our hope is that this paper inspires social scientists to apply case-control methods to other instances of extremism where data on recruits and population surveys are available. Examples include participation in the 2021 attack on the Capitol Building in Washington, DC (Pape 2021), recruitment to far-right movements and white supremacist groups (Klandermans and Nonna 2006; Simi *et al.* 2017), as well as other examples of violent extremism (della Porta 2013). It is in this spirit that we provide the `extremeR` software package so that extremism researchers working on a range of different cases can easily apply our models (see <http://extremeR.info>).

Acknowledgments. We received helpful feedback and advice from Sir David Cox, Thomas Hegghammer, Bjrn Hyland, Bent Nielsen, Jacob Aasland Ravndal, and Frank Windmeijer. Hertog, Neumann, and Maher (2021) reached out to us as we were finalizing our manuscript. Their analysis also uses leaked ISIS recruitment data to analyze the socioeconomic correlates of joining the movement. To implement the methods described in this paper, see the associated R package, as well as documentation and vignettes: <http://extremeR.info>.

Author Contributions. N.K. and C.B. conceived of the study. R.C. developed the models. R.C., C.B., and N.K. contributed to the analysis. R.C., C.B., and N.K. developed the R package and the documentation. C.B., N.K., and A.Y.Z. contributed to the data collection. C.B., N.K., and A.Y.Z. developed the literature review. All authors contributed to the writing.

Data Availability Statement. All data and code required to replicate the results and simulations described in the main article and the Supplementary Material can be found at <https://doi.org/10.7910/DVN/HYOQCD> (Cerina et al. 2023).

Supplementary Material. For supplementary material accompanying this paper, please visit <https://doi.org/10.1017/pan.2023.35>.

References

- Arab Barometer . 2014. "Arab Barometer Wave III Technical Report." Technical report, International Development Research Centre, United States Institute for Peace, University of Michigan, and Princeton University.
- Bail, C. A., F. Merhout, and P. Ding. 2018. "Using Internet Search Data to Examine the Relationship between Anti-Muslim and Pro-ISIS Sentiment in U.S. Counties." *Science Advances* 4 (6): eaao5948.
- Baio, G. 2012. *Bayesian Methods in Health Economics*. Boca Raton: CRC Press.
- Barrie, C., and N. Ketchley. 2018. "Is Protest a Safety Valve against ISIS in Tunisia?" *Washington Post*, December 10. <https://www.washingtonpost.com/news/monkey-cage/wp/2018/12/10/is-protest-a-safety-valve-against-isis-in-tunisia/>.
- Bérubé, M., R. Scrivens, V. Venkatesh, and T. Gaudette. 2019. "Converging Patterns in Pathways in and out of Violent Extremism." *Perspectives on Terrorism* 13 (6): 73–89.
- Besag, J., and C. Kooperberg. 1995. "On Conditional and Intrinsic Autoregressions." *Biometrika* 82 (4): 733–746.
- Besag, J., J. York, and A. Mollié. 1991. "Bayesian Image Restoration, with Two Applications in Spatial Statistics." *Annals of the Institute of Statistical Mathematics* 43 (1): 1–20.
- Biggs, M., and S. Knauss. 2012. "Explaining Membership in the British National Party: A Multilevel Analysis of Contact and Threat." *European Sociological Review* 28 (5): 633–646.
- Blair, G., C. C. Fair, N. Malhotra, and J. N. Shapiro. 2013. "Poverty and Support for Militant Politics: Evidence from Pakistan." *American Journal of Political Science* 57 (1): 30–48.
- Cerina, R., C. Barrie, N. Ketchley, and A. Zelin. 2023. "Replication Data for: Explaining Recruitment to Extremism: A Bayesian Hierarchical Case–Control Approach." Harvard Dataverse, V1. <https://doi.org/10.7910/DVN/HYOQCD>
- Clark, T. S., and D. A. Linzer. 2015. "Should I Use Fixed or Random Effects." *Political Science Research and Methods* 3(2):399–408.
- Corstange, D. 2009. "Sensitive Questions, Truthful Answers? Modeling the List Experiment with LISTIT." *Political Analysis* 17 (1): 45–63.
- della Porta, D. 2013. *Clandestine Political Violence*, Cambridge Studies in Contentious Politics. Cambridge: Cambridge University Press.
- Devarajan, S., et al. 2016. "Economic and Social Inclusion to Prevent Violent Extremism." Technical report, World Bank.
- Gambetta, D., and S. Hertog. 2016. *Engineers of Jihad: The Curious Connection between Violent Extremism and Education*. Princeton: Princeton University Press.
- Gelman, A., and J. Hill. 2006. *Data Analysis using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.
- Gelman, A., A. Jakulin, M. G. Pittau, and Y.-S. Su. 2008. "A Weakly Informative Default Prior Distribution for Logistic and Other Regression Models." *Annals of Applied Statistics* 2 (4): 1360–1383.
- Ghosh, J., Y. Li, and R. Mitra. 2018. "On the Use of Cauchy Prior Distributions for Bayesian Logistic Regression." *Bayesian Analysis* 13 (2): 359–383.
- Heinze, G. 2017. *Logistic Regression with Rare Events: Problems and Solutions*. Vienna: CeMSIIS-Section for Clinical Biometrics, Medical University of Vienna. <https://www.eio.upc.edu/ca/seminari/docs/georg-heinze-logistics-regression-with-rare-eventsproblems-and-solutions.pdf>.
- Heinze, G., and M. Schemper. 2002. "A Solution to the Problem of Separation in Logistic Regression." *Statistics in Medicine* 21 (16): 2409–2419.
- Hertog, S., P. Neumann, and S. Maher. 2021. "Great Expectations or Nothing to Lose? Socio-Economic Correlates of Joining the Islamic State." <https://www.belfercenter.org/event/great-expectations-or-nothing-lose-socio-economic-correlates-joining-islamic-state>.
- Jensen, M. A., A. Atwell Seate, and P. A. James. 2020. "Radicalization to Violence: A Pathway Approach to Studying Extremism." *Terrorism and Political Violence* 32 (5): 1067–1090.
- Ketchley, N., and M. Biggs. 2017. "The Educational Contexts of Islamist Activism: Elite Students and Religious Institutions in Egypt." *Mobilization* 22 (1): 57–76.
- Ketchley, N., S. Brooke, and B. Lia. 2021. "Who Supported the Early Muslim Brotherhood?" *Politics and Religion* 15 (2): 388–416.
- King, G., and L. Zeng. 2001. "Logistic Regression in Rare Events Data." *Political Analysis* 9 (2): 137–163.
- Klandermans, B., and M. Nonna. 2006. *Extreme Right Activists in Europe: Through the Magnifying Glass*. London: Routledge.
- Krueger, A. B. 2017. *What Makes a Terrorist: Economics and the Roots of Terrorism*. Princeton: Princeton University Press.
- Krueger, A. B., and J. Maleckova. 2003. "Education, Poverty and Terrorism: Is There a Causal Connection?" *Journal of Economic Perspectives* 17 (4): 119–144.

- Lancaster, T., and G. Imbens. 1996. "Case-Control Studies with Contaminated Controls." *Journal of Econometrics* 71 (1–2): 145–160.
- Mesquita, E. B. D. 2005. "The Quality of Terror." *American Journal of Political Science* 49 (3): 515–530.
- Mitts, T. 2019. "From Isolation to Radicalization: Anti-Muslim Hostility and Support for ISIS in the West." *American Political Science Review* 113 (1): 173–194.
- Morris, A. M. 2020. "Who Wants to Be a Suicide Bomber? Evidence from Islamic State Recruits." *International Studies Quarterly* 64 (2): 306–315.
- Morris, M., K. Wheeler-Martin, D. Simpson, S. J. Mooney, A. Gelman, and C. DiMaggio. 2019. "Bayesian Hierarchical Spatial Models: Implementing the Besag York mollié Model in Stan." *Spatial and Spatio-Temporal Epidemiology* 31: 100301.
- Pape, R. A. 2021. "Opinion | What an Analysis of 377 Americans Arrested or Charged in the Capitol Insurrection Tells Us." *Washington Post*, April 6.
- Riebler, A., S. H. Srbye, D. Simpson, and H. Rue. 2016. "An Intuitive Bayesian Spatial Model for Disease Mapping that Accounts for Scaling." *Statistical Methods in Medical Research* 25 (4): 1145–1165.
- Robinson, W. S. 1950. "Ecological Correlations and the Behavior of Individuals." *American Sociological Review* 15 (3): 351–357.
- Rosenfeld, B. 2017. "Reevaluating the Middle-Class Protest Paradigm: A Case-Control Study of Democratic Protest Coalitions in Russia." *American Political Science Review* 111 (4): 637–652.
- Rosenfeld, B. 2018. "A Case-Control Method for Studying Protest Participation and Other Rare Events: Application to Ukraine's Euromaidan." Paper presented at the 2018 Annual Conference of the American Political Science Association, 1–41.
- Rota, C. T., J. J. Millsbaugh, D. C. Kesler, C. P. Lehman, M. A. Rumble, and C. M. Jachowski. 2013. "A Re-Evaluation of a Case-Control Model with Contaminated Controls for Resource Selection Studies." *Journal of Animal Ecology* 82 (6): 1165–1173.
- Selb, P., and S. Munzert. 2011. "Estimating Constituency Preferences from Sparse Survey Data Using Auxiliary Geographic Information." *Political Analysis* 19 (4): 455–470.
- Simi, P., K. Blee, M. DeMichele, and S. Windisch. 2017. "Addicted to Hate: Identity Residual among Former White Supremacists." *American Sociological Review* 82 (6): 1167–1187.
- Simpson, D., H. Rue, A. Riebler, T. G. Martins, S. H. Srbye. 2017. "Penalising Model Component Complexity: A Principled, Practical Approach to Constructing Priors." *Statistical Science* 32 (1): 1–28.
- Skare, E. 2022. "Affluent and Well-Educated? Analyzing the Socioeconomic Backgrounds of Fallen Palestinian Islamist Militants." *Middle East Journal* 76 (1): 72–92.
- Sterman, D., and N. Rosenblatt. 2018. "All Jihad Is Local: ISIS in North Africa and the Arabian Peninsula." Technical report II, New America.
- Zelin, A. Y. 2018. "Tunisian Foreign Fighters in Iraq and Syria." Technical report, Washington Institute for Near East Policy.