Title
Nonrandom Timing, Biased Inference? Rethinking Survey-Based Discontinuities

ISBN
9798252408187

Author
Stecher, Elayne

Publication Date
2026-06-12

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Nonrandom Timing, Biased Inference?
Rethinking Survey-Based Discontinuities

A thesis submitted in partial satisfaction
of the requirements for the degree
Master of Science in Statistics

by

Elayne Stecher

2026

ABSTRACT OF THE THESIS

Nonrandom Timing, Biased Inference?
Rethinking Survey-Based Discontinuities

by

Elayne Stecher
Master of Science in Statistics
University of California, Los Angeles, 2026
Professor Chad Hazlett, Chair

Unexpected events during survey fieldwork are often used to estimate causal effects by comparing respondents interviewed before and after the event. This thesis argues that the credibility of these designs depends not only on whether the event was unexpected, but also on whether fieldwork rollout shaped the composition of respondents observed before and after the event. When rollout is geographically structured, the event cutoff may divide respondents by place and composition in ways related to potential outcomes. I examine this problem using survey data around the 2015 Garissa University College attack in Kenya. The case shows that standard design repairs, including restricting to common support and adjusting for covariates, improve comparability but do not necessarily eliminate rollout-related confounding. Except where stronger assumptions can be justified, unexpected-event survey designs belong in the world of selection on observables. Their advantage is not automatic as-if random timing, but diagnostic transparency: researchers can observe the fieldwork sequence, assess overlap, restrict claims to supported comparisons, and test whether similar estimates appear at non-event cutoffs.

The thesis of Elayne Stecher is approved.

Ian Lundberg

Frederic Schoenberg

Chad Hazlett, Committee Chair

University of California, Los Angeles

2026

To Francesco,

who first encouraged me to learn to program

and then to pursue this degree:

`print("this one's for you")`.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my colleagues in the Departments of Statistics and Political Science for their camaraderie, and my Chair, Chad Hazlett, for his guidance, mentorship, and encouragement.

CHAPTER 1

Introduction

Over the past two decades, scholars have increasingly leveraged major events that occur during survey fieldwork to estimate causal effects on political attitudes, beliefs, and behavior. This approach, often referred to as Unexpected Event during Survey Design (UESD), compares respondents interviewed before a focal event to respondents interviewed afterward. Because events such as terrorist attacks, natural disasters, political scandals, or sudden policy announcements are often unforeseen, the design is typically framed as a natural experiment in which interview timing generates as-if random exposure to treatment.

In many applied UESD studies, the credibility of the design is motivated by characteristics of the event: that it was sudden, unanticipated, salient, and plausibly unrelated to the survey itself. Yet, recent methodological work has shown that these conditions are not sufficient. Muñoz et al. (2020) show that credible UESD inference requires researchers to justify the comparability of pre- and post-event respondents, not merely the unexpectedness of the event. Bertoli et al. (2026) further formalize sources of bias in event-through-survey designs and propose solutions to improve pre/post comparability. This thesis builds on that work by focusing particularly on spatial rollout during fieldwork as a central threat to inference in event-through-survey designs.

So, what can researchers do? When simple as-if random timing is not credible, UESD designs should be understood less as natural experiments and more as selection-on-observables designs with unusually transparent assignment processes. Researchers can often observe the fieldwork sequence, diagnose geographic and compositional shifts, assess common support, adjust for observed outcome-relevant covariates, and use placebo cutoffs to test whether similar apparent effects appear at non-event dates. These diagnostics do not prove causal identification, but they make the assumptions and scope of the comparison more visible.

I examine these issues using Ipsos survey data from Kenya collected around the 2015 Garissa University College attack. The case illustrates how spatially structured fieldwork can affect the comparison between respondents interviewed before and after an unexpected event. I use the case to show how existing UESD diagnostics can help determine not only whether pre- and post-event respondents are balanced, but also which comparisons the data can credibly sustain once survey rollout is taken into account.

CHAPTER 2

Literature Review

This chapter reviews literatures relevant to event-based pre/post survey designs. I first discuss UESD as a quasi-experimental strategy, then review methodological critiques of its identifying assumptions, and finally connect these critiques to selection-on-observables, common support, and diagnostic practice. This review motivates Chapter 3: when fieldwork rollout is spatially structured, UESD should be understood as a conditional design rather than as-if random assignment by default. Its credibility depends on whether the realized fieldwork process shapes pre/post composition in ways related to respondents' potential outcomes.

2.1 UESD as a Quasi-Experimental Design

Unexpected Event during Survey Design (UESD) studies have become increasingly popular because they appear to offer a rare opportunity to estimate causal effects using observational survey data. When an unexpected event occurs during fieldwork, researchers compare respondents interviewed before the event to respondents interviewed afterward, treating interview timing as a source of quasi-random exposure. Like other natural-experiment designs, UESD seeks to approximate random assignment without researcher-controlled randomization.

UESD approaches not only seem to provide causal leverage, they also have several attractive practical features: they allow researchers to use already-gathered data to answer new research questions; they tend to focus on real-world shocks with greater external validity than researcher-created interventions; and they allow scholars to study phenomena that would be unethical or otherwise difficult to induce experimentally, including violent or socially significant events. Applied UESD studies use events such as terrorist attacks (e.g., Dinesen and Jæger, 2013; Muñoz et al., 2020), protests (e.g., Sangnier and Zylberberg, 2017; Frye and Borisova, 2019), police violence (e.g., White et al., 2018; Curtice, 2021; see Nägel and Nivette, 2023), political scandals (e.g., Ares and Hernández, 2017), and public health crises (e.g., Sibley et al., 2020) to study short-term changes in attitudes, opinions or behaviors.

The topical breadth of these applications also speaks to the growing popularity of UESD in contemporary research. In their review of papers across the social sciences, Muñoz et al. (2020) identify 44 English-language UESD studies published between 1969 and 2019 in political science, sociology, economics, and related journals. Looking at just three top political science journals between 2015–2025, Bertoli et al. (2026) document 37 event-based pre/post survey designs, about half of which use unexpected events as their key identification leverage. In criminology alone, Nägel and Nivette (2023) identify 12 studies using unexpected events to study trust in police, half of which were published between 2019 and 2021; the oldest was published in just 2012. These reviews demonstrate that UESD has moved from an opportunistic empirical strategy to a recognizable design tradition, particularly in political science and other social scientific domains.

At the same time, the design's growing popularity has not always been matched by attention to the assumptions required for credible causal inference. In many applications, the unexpectedness and salience of the event do substantial work in motivating the identification claim. But event unexpectedness does not by itself establish that respondents interviewed before and after the event are comparable.

2.2 Methodological Critiques of UESD

2.2.1 The Identification Challenge

The key methodological question is whether respondents interviewed before and after the event are comparable in potential-outcomes terms. Would respondents interviewed after the event have had similar outcomes to respondents interviewed before the event if the event had not occurred?

Let $D_i = 1$ indicate that respondent i was interviewed after the event and $D_i = 0$ indicate that the respondent was interviewed before the event. Let $Y_i(1)$ and $Y_i(0)$ denote respondent i's potential outcomes under post-event and pre-event interview status, respectively. The observed outcome is

$$Y_i = D_i Y_i(1) + (1 - D_i) Y_i(0),$$

assuming consistency: the observed response corresponds to the potential outcome under the respondent's realized pre/post interview status. Under this notation, the key identifying condition for a simple pre/post comparison is that post-event interview status is independent of potential outcomes:

$$(Y_i(1), Y_i(0)) \perp D_i.$$

This condition is often referred to as ignorability: pre/post interview status must be unrelated to the potential outcomes respondents would express under either condition. In words, respondents interviewed after the event must be comparable to those interviewed before the event in terms of what their outcomes would have been under either pre- or post-event interview status.

2.2.2 Existing Guidance and Remaining Gaps

Recent methodological work on UESD starts from the premise that event unexpectedness alone is not enough for causal identification. These studies argue that researchers must do two things: assess whether the pre/post comparison is plausibly comparable, and state clearly what assumptions remain after any diagnostics, restrictions, or adjustments are applied.

Muñoz et al. (2020) provide the foundational methodological critique of UESD as a causal design. They emphasize that UESD typically uses interview timing as the basis for comparing respondents interviewed before and after an event, and that this comparison is credible only under strong assumptions about pre/post comparability, event excludability, and exposure or compliance. In their review of applied UESD studies, they show that many papers rely heavily on the event's unexpectedness, salience, and apparent independence from survey operations, while using recommended practices such as placebo tests, nonresponse analyses, and checks related to exposure or compliance in fewer than 30% of studies. Muñoz et al. therefore recommend that researchers consider covariate balance tests, bandwidth restrictions, covariate adjustment or matching, placebo analyses, falsification outcomes, nonresponse checks, and attention to treatment compliance. They also caution that many UESD estimates should be interpreted as intent-to-treat effects of post-event interview status rather than direct effects of exposure to the event.

Similarly, in their systematic review of UESD studies on trust in police, Nägel and Nivette (2023) show that applications vary substantially in how rigorously they evaluate identification assumptions. They find that checks related to pre/post comparability are more common than checks assessing whether other time-varying shocks coincide with the event. This pattern suggests that applied UESD research often recognizes the need to evaluate comparability across the cutoff, but also echoes the concern in Muñoz et al. (2020) that studies vary in how completely they assess threats to identification.

Bertoli et al. (2026) extend this conversation by formalizing sources of bias in event-through-survey designs, including differences between pre- and post-event respondents, temporal factors, anticipation, and differential misreporting. They also show that design choices intended to improve comparability involve tradeoffs: quota sampling and rolling cross-sections may reduce some pre/post imbalance without eliminating it, while panels hold respondents constant across waves but can introduce panel conditioning. Their proposed dual randomized survey is therefore a prospective design-stage solution: it can improve pre/post comparability when researchers can plan data collection around an anticipated event, but it offers less guidance for retrospective analyses of already-fielded surveys.

Taken together, these studies shift the methodological conversation away from the simple claim that unexpected events automatically create credible causal designs. They emphasize that researchers must defend pre/post comparability, distinguish post-event interview status from actual exposure, assess collateral events and time trends, and state what assumptions remain after restrictions or adjustments. Recommended checks are useful because they make these assumptions more explicit rather than leaving the causal claim to rest on unexpectedness alone.

This thesis builds on that guidance by arguing that the threat of rollout-related confounding deserves more central attention. The aforementioned methodological work recognizes fieldwork organization as one possible threat to UESD, alongside anticipation, collateral events, nonresponse, exposure, and time trends. I focus on fieldwork rollout because, in face-to-face surveys, the realized movement of survey teams through space can directly shape the composition of respondents observed before and after the event. Balance tests can show whether observed respondent characteristics differ across pre/post status, but they do not reveal whether those differences are products of the realized fieldwork sequence, nor can they rule out imbalance in unobserved factors. Without specific attention to this concern, scholars may incorrectly treat reasonable balance on observed covariates as sufficient for causal identification.

The next section therefore turns to survey implementation, focusing on how fieldwork timing, rollout, and geographic sequencing can shape the composition of pre- and post-event respondents in ways that undermine comparability across the cutoff.

2.3 Survey Implementation, Fieldwork Timing, and Balance

This section focuses on how survey implementation can generate pre/post imbalance. Respondents fall into pre- or post-event status because their interviews occur before or after the event cutoff, but interview timing is shaped by fieldwork schedules, interviewer routes, quotas, recontacting policies, respondent reachability, safety conditions, and administrative constraints. If the event changes sampling procedures or response patterns, pre/post comparisons may reflect changes in survey implementation rather than the event's effect on outcomes.

However, even when no post-event fieldwork policy change occurs, preexisting fieldwork rollout can still produce geographically structured differences between pre- and post-event respondents. Survey firms rarely distribute interviews randomly across all locations and respondent types each day of fieldwork. This is especially true in face-to-face surveys, where enumerators often work through clusters, neighborhoods, municipalities, or regions under transportation, staffing, or safety constraints. In these settings, the event cutoff intersects with the preexisting fieldwork schedule. Respondents interviewed before and after the event may therefore differ not because the event changed survey operations, but because survey operations were already moving through the sample in a systematic way.

As such, observed imbalances between pre- and post-event respondents are not simply a technical nuisance to be resolved with reweighting or covariate adjustment. They provide evidence about whether the realized fieldwork sequence may have placed different kinds of respondents on either side of the event cutoff in ways related to their untreated potential outcomes. If geography, urbanicity, ethnicity, socioeconomic status, or security exposure predicts both interview timing and the potential outcomes, then the ignorability assumption is threatened.

Addressing this threat requires researchers to defend a conditional comparability claim rather than relying on as-if random timing: that, within comparable values of the covariates that structure fieldwork timing and predict the outcome, interview timing is unrelated to potential outcomes. Whether that claim is even plausible depends on whether the data contain comparable pre- and post-event respondents once those fieldwork- and geography-related differences are taken into account, an issue I turn to in the next section.

2.4 Selection-on-Observables, Common Support, and Diagnostics

When fieldwork rollout structures interview timing, the central concern is not observed imbalance itself, but imbalance in potential outcomes. Respondents interviewed before and after the event may have differed in their untreated outcomes even if the event had never occurred. Observable covariate imbalance is useful because it can help diagnose whether the fieldwork process sorted respondents in ways related to their potential outcomes, but balance on observed covariates does not prove ignorability.

In this setting, the identifying assumption is best understood as a conditional ignorability claim. Formally, the assumption becomes $(Y_i(1), Y_i(0)) \perp D_i \mid X_i$, where X_i denotes observed covariates used to define comparable pre- and post-event respondents. In the broader causal inference literature, this is a selection-on-observables or unconfoundedness assumption: conditional on observed covariates, assignment to post-event status is assumed to be independent of potential outcomes (Rosenbaum and Rubin, 1983; Imbens and Rubin, 2015). Diagnostics and adjustments therefore play a limited but important role. Balance checks, maps, covariate adjustment, weighting, and placebo cutoffs do not create causal identification on their own. They help assess whether the fieldwork produced pre- and post-event groups that differ in ways related to potential outcomes, and they clarify what comparison the data can support.

For conditional comparisons to be credible, researchers also need overlap or common support: comparable pre- and post-event units must exist across the covariate values used to define the comparison (Imbens and Rubin, 2015). Formally, this requires that for the relevant values of X_i, there must be respondents observed both before and after the event. Equivalently, $0 < P(D_i = 1 \mid X_i) < 1$ for the covariate values used in the comparison. This concern matters even when the event is unanticipated, because fieldwork rollout may place some regions, respondent types, or covariate profiles entirely on one side of the cutoff. When common support is lacking, researchers should either restrict the analysis to the region of overlap or characterize the estimate as applying only to those respondents for whom comparable pre- and post-event observations exist. Lack of common support can be detected through cross-tabulations for discrete covariates, maps of pre/post coverage by geography, propensity score overlap plots, and inspection of covariate distributions by treatment status. Chapter 3 applies these formal conditions to spatial fieldwork rollout, showing what conditional ignorability would require and how geographically structured rollout can undermine it.

CHAPTER 3

Theory: Fieldwork Rollout and Assignment

3.1 Event Timing and Fieldwork Rollout

Treatment status is defined by whether respondent i was interviewed before or after the event. I write this as:

$$D_i = \mathbf{1}(t_i > T)$$

where t_i is respondent i's survey date and T is the event date. In many applications, surveys from $t_i = T$ are excluded from analysis, as it is difficult to ascertain whether or not individual i has been aware of the focal event prior to completing the survey. Under the strongest version of the UESD claim, the event date T is unexpected and plausibly exogenous to the survey process, so whether t_i falls before or after T is treated as-if random. A key contribution of this thesis is to ask not only whether T is exogenous, but also how t_i is generated. Because D_i is mechanically determined by t_i, any systematic structure in interview timing can become systematic structure in the composition of pre- and post-event respondents.

Fieldwork rollout can therefore be understood as the process that orders respondents in time. In an idealized UESD, this ordering would be unrelated to respondents' potential outcomes. In practice, however, the ordering of respondents is often produced by the survey organization's route through the sample. The event date then cuts through that ordering. Thus, even if the event date T is unexpected, whether respondent i appears before or after the cutoff depends on where the fieldwork process was in the interview sequence when the event occurred. Put otherwise, even if T is truly unanticipated and exogenous, it does not follow that D_i is as-if random.

3.2 Rollout-Related Confounding and Geography

In face-to-face UESD, geography is often the most visible sign of rollout-related confounding. The concern is not that geography itself is always unobserved, as some geographic features may be measured and included in X_i. The deeper concern is that geographic rollout may proxy for unobserved or incompletely observed place-based factors that are related both to pre/post interview status and to potential outcomes.

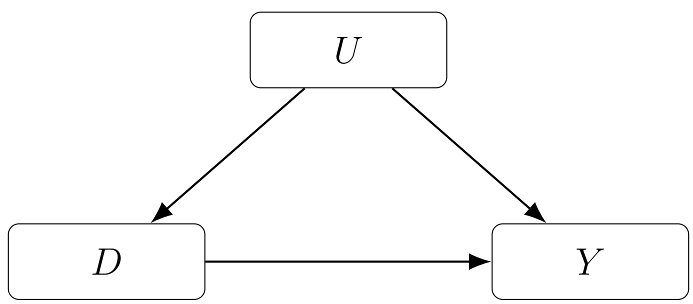

Figure 3.1: Rollout-related confounding in UESD. Unobserved or incompletely observed rollout-related factors (U) may be related to both pre/post interview status (D) and the outcome (Y).

I use U to denote unobserved or incompletely observed place-based and rollout-related characteristics that could affect both whether respondents are observed before or after the event and the potential outcomes respondents would express under each condition. These characteristics may include local enumerator decisions, accessibility, security conditions, unmeasured political context, or other features of places that are not fully captured in the data. Measured geographic variables, such as province, urbanicity, or distance from the event, are better understood as observed proxies or components of this broader confounding structure and can be included in X_i.

The DAG in Figure 3.1 therefore represents the confounding structure that motivates concern, not a claim that all relevant features of geography are unobserved. Conditioning on observed X_i can address measured differences related to rollout, but it cannot guarantee that the unobserved components of U are balanced. The design therefore moves toward a conditional ignorability claim: researchers must argue that, after conditioning on observed covariates, remaining unobserved rollout-related differences are not related to potential outcomes.

3.3 Bias: Imbalance $\times$ Impact

The DAG in Figure 3.1 identifies the backdoor path from D to Y through U. Rollout-induced confounding is most concerning when unobserved or incompletely observed characteristics both differ across pre- and post-event respondents and are related to the outcomes those respondents would have expressed under either condition. I describe this as an imbalance $\times$ impact problem. Imbalance captures whether a confounder differs across pre/post status, while impact captures whether that confounder predicts potential outcomes.

For a binary confounder U, imbalance can be written as $E[U_i \mid D_i = 1] - E[U_i \mid D_i = 0]$. Impact refers to the relationship between U and potential outcomes, $Y_i(0)$ and $Y_i(1)$. Bias is most concerning when both quantities are large in magnitude. A factor that strongly predicts potential outcomes but is balanced across pre- and post-event status is less concerning; a factor that is highly imbalanced but unrelated to potential outcomes is similarly less consequential. In this sense, the imbalance $\times$ impact framework restates the ignorability concern from Chapter 2: pre/post status becomes problematic when it is related to the potential outcomes respondents would have expressed under either condition.

Because U is unobserved or only partially observed, researchers cannot directly assess either the $U \to D$ path or the $U \to Y$ path. Instead, the empirical checks examine observed covariates X_i. Balance checks ask whether observed covariates differ acriss pre/post status, or $X_i \to D_i$. Outcome-prediction exercises ask whether observed covariates predict outcomes, or $X_i \to Y_i(0)$ in the pre-event period and $X_i \to Y_i(1)$ in the post-event period. These diagnostics do not prove that the relevant unobserved confounders are balanced or irrelevant. Rather, they show whether observed features of the fieldwork process and respondent composition are imbalanced and outcome-relevant, which can signal concern about similar unobserved rollout-related characteristics and guide the covariates used in the selection-on-observables adjustment below.

3.4 From As-If Random Timing to SOO+

When fieldwork rollout may shape pre/post respondent composition in ways related to potential outcomes, the identifying claim shifts from unconditional comparability to conditional comparability. Rather than assuming $(Y_i(1), Y_i(0)) \perp D_i$, researchers must defend the claim that $(Y_i(1), Y_i(0)) \perp D_i \mid X_i$, where X_i includes observed covariates related to fieldwork timing and potential outcomes. This is a selection-on-observables claim.

I characterize this version of UESD as SOO+: it relies on conditional comparability, like ordinary selection-on-observables, but retains additional diagnostic leverage because the event cutoff is known, the fieldwork sequence is often observable, and placebo cutoffs in time can be used to assess whether estimated effects are tied to the actual event.

Following the distinction in Muñoz et al. (2020), the estimand should generally be understood as the effect of post-event interview status rather than the direct effect of exposure to the event. In the full survey sample, this post-status estimand can be written as $\tau = E[Y_i(1) - Y_i(0)]$. Because post-event status may not correspond perfectly to actual exposure or event awareness, this quantity is best interpreted as an intent-to-treat effect of being observed after the event cutoff.

If researchers restrict the analysis to a supported subset S, the estimand changes: it applies to that subset rather than to the full survey sample. The supported-sample estimand is $\tau_S = E[Y_i(1) - Y_i(0) \mid i \in S]$, where S is the subset of respondents for whom comparable pre- and post-event observations exist. This restriction is a consequence of the common support problem: it does not solve rollout-induced confounding by itself, but it clarifies the population for which a conditional comparison is being defended.

3.5 Placebo Cutoffs

Placebo cutoffs in time are an important part of the "plus" in SOO+. In ordinary selection-on-observables designs, researchers often lack a natural way to ask whether the same adjustment strategy would generate an apparent effect in the absence of a treatment or event. In UESD, the fieldwork calendar can sometimes provide such placebo comparisons. If the same adjusted pre/post specification produces sizable differences at non-event cutoffs, this suggests that residual fieldwork composition may still be generating apparent effects. Placebo cutoffs therefore do not establish causality, but they help assess whether the estimated change is distinctive to the actual event date or instead reflects broader temporal or rollout-related shifts in the survey sample.

To be useful, a placebo cutoff should not mix observations from before and after the true event. If the true event falls inside a placebo comparison window, the placebo estimate may partially capture the real event comparison rather than only fieldwork-related composition shifts. I therefore restrict placebo windows so that both sides of each placebo cutoff remain entirely on the same side of the true Garissa cutoff.

For a placebo cutoff T^*, define $D_i^* = \mathbf{1}(t_i > T^*)$, where T^* is an alternative event date and $T^* \neq T$. For each T^*, estimate the same observed pre/post contrast using D_i^* as the placebo event indicator. Because no focal event occurs at T^*, the resulting estimate is not a causal treatment effect. Instead, the placebo cutoff analysis asks whether similar apparent effects emerge when the event date is moved to dates when no event occurred.

If placebo cutoffs generate estimates similar in magnitude to the estimate at the true event cutoff, the main estimate may reflect fieldwork-induced compositional change or time trends rather than the causal effect of the event. This logic is especially useful for UESD because nearby placebo cutoffs preserve much of the same fieldwork process. If the true cutoff estimate is unusually large relative to the placebo estimates, the event-based interpretation becomes more credible.

CHAPTER 4

Empirical Setting and Data

The preceding chapter developed a framework for understanding how fieldwork rollout can shape the composition of respondents observed before and after an unexpected event. I now apply this framework to survey data from Kenya collected around the April 2, 2015 Garissa University College attack. The case is useful because it combines a plausibly unexpected and salient security shock with survey fieldwork whose timing and geography can be examined directly.

4.1 Kenya: Security and Political Context

Kenya is an East African country on the Indian Ocean bordering Tanzania, Uganda, South Sudan, Ethiopia, and Somalia. It is home to more than 40 ethnic groups and is one of the most populous countries in Africa. Although Kenya is often considered a relatively stable regional hub, its proximity to conflict in Somalia has created persistent security challenges. These challenges became especially salient with the rise of Harakat Al-Shabaab Al Mujahideen, more commonly referred to as Al-Shabaab.

Al-Shabaab emerged in the context of Somalia's prolonged state collapse after the fall of Mohamed Siad Barre in 1991. Initially a youth militia movement, Al-Shabaab rose to prominence through its resistance to the U.S.-backed Ethiopian military intervention in Somalia in the mid- to late 2000s. This resistance to foreign powers lent Al-Shabaab a degree of legitimacy within Somalia, where the group also provides social services and elements of governance, particularly in the south of the country (Mwangi, 2012). By the 2010s, Al-Shabaab had become one of the most significant armed actors in the region, combining territorial control, cross-border violence, and propaganda to recruit local and foreign fighters (Mwangi, 2012; Speckhard and Shajkovci, 2019).

Al-Shabaab began launching attacks in Kenya before the 2010s, but the frequency and intensity of those attacks increased after the Kenyan military entered Somalia. Beginning in 2011, the Kenyan Defence Forces (KDF) conducted counterterrorism operations in southern Somalia, first through Operation Linda Nchi and subsequently as part of the African Union Mission in Somalia (AMISOM). The stated goal of Operation Linda Nchi was to secure the Kenya-Somalia border and weaken Al-Shabaab. East African security analysts described the operation as a major shift in Kenya's security posture and warned that intervention in Somalia could invite retaliation inside Kenya (Atta-Asamoah and Kisiangani, 2011; Miyandazi, 2012). Anderson and McKnight (2015) similarly describe Kenya's intervention as a turning point that contributed to a prolonged conflict with Al-Shabaab and subsequent "blowback" through attacks inside Kenya.

These operations made counterterrorism and the deployment of military forces in Somalia salient domestic political issues within Kenya. KDF activity in Somalia was not simply a foreign-policy question, but one tied to domestic debates over national security, border control, state capacity, and the costs of military intervention. This security context also intersected with Kenya's domestic political cleavages: Kenyan electoral politics during this period were organized in significant part around ethnic-coalition alignment. Ethnicity strongly predicted vote choice in the 2013 election, especially among groups with co-ethnic presidential or deputy presidential candidates (Ferree et al., 2014). These political cleavages matter because they imply that reactions to a violent shock might differ along ethnic- or political coalition-based lines.

4.2 The Garissa University College Attack

On April 2, 2015, gunmen affiliated with Al-Shabaab attacked Garissa University College in northeastern Kenya. The attackers entered the campus in the early morning, targeted students, and held parts of the university for several hours before Kenyan security forces ended the siege. The attack killed 148 people, most of them students, and became one of the deadliest terrorist attacks in Kenya's history. Al-Shabaab claimed responsibility and framed the attack as retaliation for Kenya's military presence in Somalia.

The attack received extensive national and international attention. It also generated immediate debate over the Kenyan state's counterterrorism strategy, the capacity of security forces to prevent attacks, and the costs of the KDF presence in Somalia. Because the attack occurred on a specific date and was highly salient, it provides a clear event cutoff for an event-based pre/post design.

4.3 Ipsos Kenya Survey Fieldwork

The empirical analysis uses Ipsos survey data collected in Kenya during the period surrounding the Garissa University College attack. The survey was fielded face-to-face, which makes it especially useful for examining the relationship between interview timing and geographic rollout. Because the data include interview dates and location, they allow me to observe not only who was interviewed before and after the attack, but also where fieldwork was taking place over time.

Fieldwork began on March 28, 2015 and ended on April 7, 2015. The Garissa attack occurred on April 2, 2015, during the survey period. The available data do not fully reconstruct Ipsos's fieldwork protocol, including enumerator route assignments or callback rules. However, the data do contain enough information to characterize realized fieldwork rollout: interview date, respondent location, and relevant demographic/geographic covariates.

I therefore focus on the realized sequencing of interviews rather than on Ipsos's intended fieldwork plan.

4.4 Measures: Treatment and Outcomes

The treatment indicator is post-event interview status: respondents interviewed before April 2, 2015 are coded as pre-event, and respondents interviewed after April 2 are coded as post-event. Following convention, respondents interviewed on the day of the attack are excluded because it is conceivable, although improbable in this case, that respondents on April 2 could have been surveyed before learning of the attack.

The primary outcome is support for the KDF mission in Somalia, coded as 1 for respondents who approve of the mission and 0 for all other non-missing responses. This outcome is substantively appropriate because the Garissa attack was publicly linked to Al-Shabaab and Kenya's military intervention in Somalia. It allows me to examine whether support for counterterrorism policy changed after a salient terrorist attack, while assessing how much that comparison depends on fieldwork rollout.

4.5 Why This Case?

This case is useful because it contains the core characteristics that make UESD appealing in the first place. The Garissa University College attack was sudden, highly salient, and plausibly capable of shifting attitudes toward counterterrorism policy and the KDF mission in Somalia. It also occurred while a national survey was already in the field, allowing for a straightforward pre/post comparison between respondents interviewed before and after the attack.

At the same time, the case also illustrates the central concern of this thesis. The survey was conducted through face-to-face fieldwork, meaning that interview timing was tied to the movement of survey teams through space. If teams were concentrated in different places before and after the attack, then pre/post respondents may differ in potential outcomes for reasons related to fieldwork rollout rather than the attack itself.

The goal is not to show that all UESD applications suffer from the same problem, or that the Kenya estimate generalizes to other settings. Rather, the goal is to show how an apparently credible event-based pre/post design can depend on the realized fieldwork sequence that shapes which respondents are observed on either side of the cutoff. Maps, balance checks, common-support diagnostics, outcome-predictive covariates, and placebo cutoffs can then clarify whether the comparison is credible and how narrowly its conclusions should be scoped.

CHAPTER 5

Analysis and Discussion

5.1 Spatial Diagnostics and Limits of RD Interpretation

5.1.1 Fieldwork Rollout

There are two distinct fieldwork concerns with the 2015 Ipsos Kenya survey data. The first is event-induced change in fieldwork: if the attack caused Ipsos to alter where or how interviews were conducted, then post-attack respondents may differ from pre-attack respondents because the survey process changed after the event. The second is preexisting rollout: even if the attack did not alter fieldwork elsewhere, the survey may already have been moving through Kenya in a spatially patterned way.

I am most concerned about event-induced fieldwork changes in North Eastern Province, where the attack occurred. Such changes would represent a direct shift in the sampling process caused by the event, rather than the preexisting rollout problem emphasized here. I therefore show the full fieldwork file in the spatial checks below, but exclude North Eastern Province from the main adjusted analyses.

However, excluding this province does not eliminate the broader identification problem. Figure 5.1 maps survey interview locations by fieldwork date. It uses the full fieldwork file, including April 2 interviews and North Eastern Province observations, to show the realized spatial distribution of interviews by date. This version is useful because it shows both forms of fieldwork concern: the direct event-induced concern in North Eastern Province and the broader spatial sequencing of survey rollout across the rest of the country.

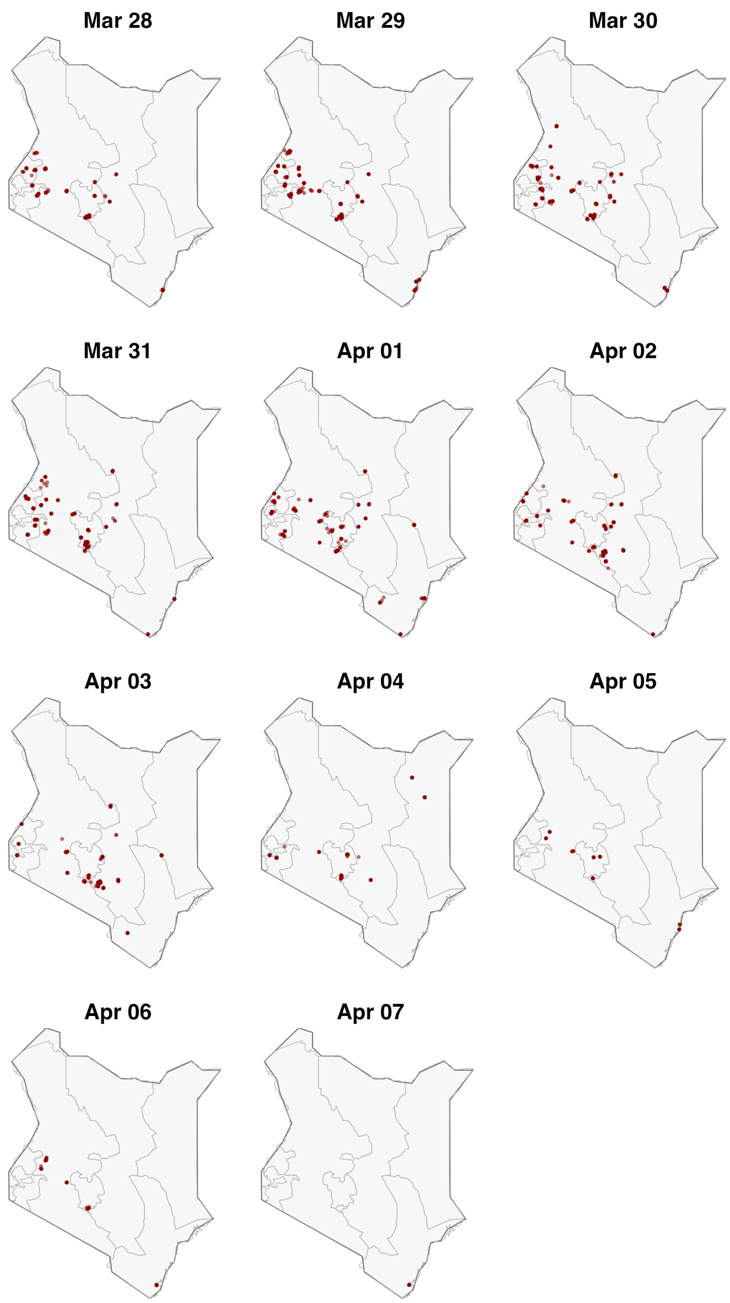

Figure 5.1: Survey interview locations by fieldwork date.

Figure 5.1 shows that interviews were not distributed randomly across the country each day. Instead, the realized sample shifts geographically over time, with survey teams concentrated in different places on different dates. This is precisely the kind of fieldwork structure that can make pre/post status related to respondent composition, even when the event is unexpected and survey procedures outside North Eastern did not change in direct response to the attack. The maps therefore suggest that respondents' pre/post status may reflect both the event cutoff and the geography of survey rollout.

Appendix Figure A.1 shows the number of interviews completed by date. While there is no clear collapse in interview volume immediately after the attack, there is a general downward trend in surveys per day after March 29.

5.1.2 Limits of Regression Discontinuity

One might argue that a regression discontinuity (RD) design around April 2 would address some of the concerns raised by a simple pre/post comparison. If respondents interviewed just before and just after the attack were otherwise comparable, restricting attention to a narrow window around the cutoff could make the comparison more credible.

In this case, however, the spatial rollout patterns make the continuity assumption difficult to defend. An RD interpretation would require potential outcomes to evolve smoothly across the cutoff. But when face-to-face survey teams move through locations in a clustered or spatially organized way, adjacent interview dates may contain respondents from different places and populations. Temporal proximity to the cutoff therefore does not necessarily imply comparability in potential outcomes.

The maps in Figure 5.1 suggest that interview timing was tied to spatial rollout. Respondents interviewed immediately before and after April 2 may differ not only in exposure to the Garissa attack, but also in the composition and location of respondents included in the survey on those dates. Narrowing the temporal bandwidth may therefore fail to solve the identification problem, and may even intensify it if the days closest to the event correspond to different fieldwork locations.

I explore this concern by examining support for the KDF across the fieldwork period. Figure 5.2 shows daily mean support for KDF operations. The purpose of this figure is not to estimate the effect of the attack, but to assess whether the observed outcome evolves smoothly over interview dates. Mean support varies noticeably from day to day, including within periods that do not cross the Garissa cutoff. Combined with the spatial rollout shown in Figure 5.1, this pattern raises concerns about interpreting a temporal discontinuity as the effect of the attack.

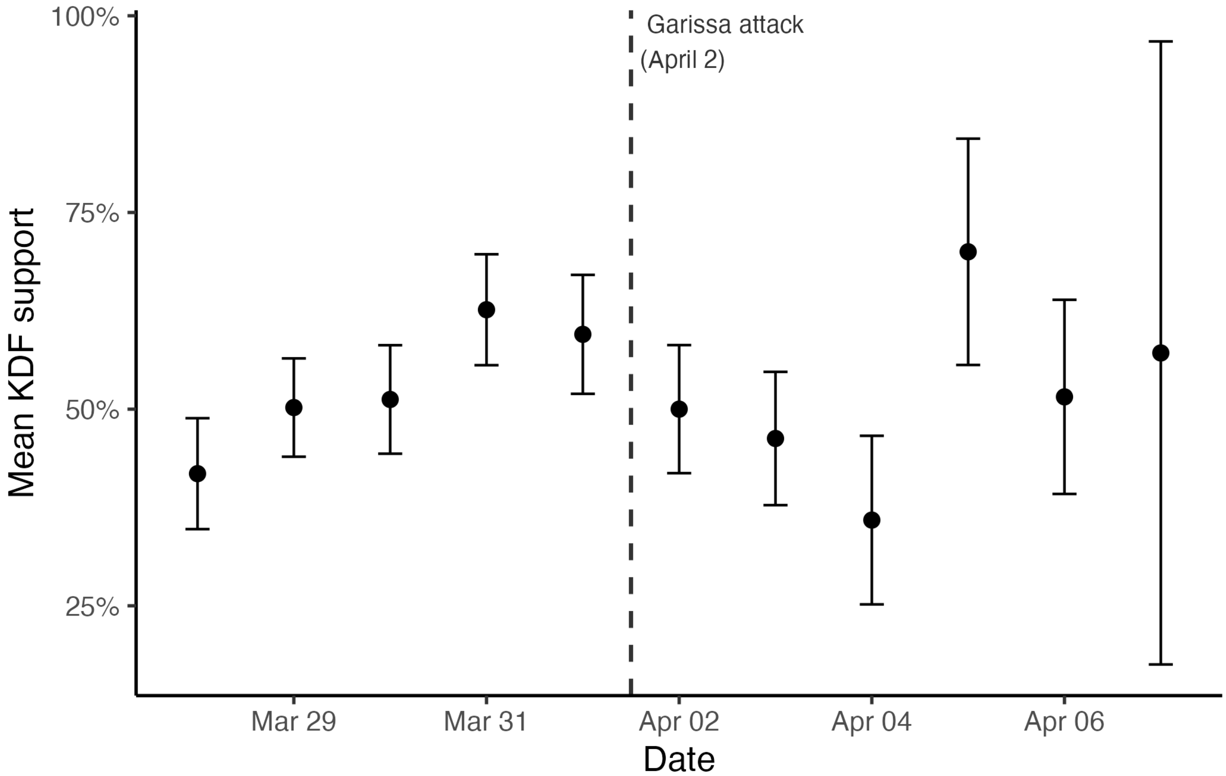

Figure 5.2: Mean support for the KDF by day.

For these reasons, I do not treat RD as the primary design. The running variable is coarse, interview timing is measured in days over a short fieldwork window, there are few support points around the cutoff, and treatment timing on the attack date is ambiguous. More importantly, the day-to-day changes in respondent composition and observed KDF support make the continuity assumption difficult to defend.

5.2 Naive Pre/Post Estimates

I next estimate the naive pre/post comparison that a simple UESD design would use as its starting point. This estimate compares respondents interviewed before the Garissa attack to respondents interviewed afterward, without adjusting for geography, respondent composition, or fieldwork rollout.

Let D_i indicate whether respondent i was interviewed after the attack. The naive estimate is:

$$\hat{\tau}_{naive} = E[Y_i \mid D_i = 1] - E[Y_i \mid D_i = 0],$$

where Y_i is support for KDF operations.

Table 5.1: Naive pre/post estimate of KDF support

Pre mean	Post mean	Difference (Post - Pre)	SE	p-value	N pre	N post
0.526	0.48	-0.047	0.032	0.147	984	323

Table 5.1 reports the naive pre/post estimate using the full provincial coverage, excluding surveys conducted on April 2 following convention. Mean support for KDF operations was 52.6 percent before the Garissa attack and 48.0 percent after the attack, a decline of 4.7 percentage points. This difference is not statistically distinguishable from zero at conventional levels ($p = 0.147$). As a benchmark, the pooled naive estimate therefore suggests a modest decline in KDF support after the attack, but the spatial diagnostics above indicate that this comparison should not be interpreted causally without further adjustment.

5.3 Balance and Common Support

5.3.1 Full Sample Balance

I next examine whether respondents interviewed before and after the attack were drawn from different places or respondent groups. This balance check is diagnostic: it does not determine the final analytic sample and it does not prove bias. Instead, it shows where pre/post respondent composition is most visibly entangled with fieldwork geography. I examine geography and rollout-related variables, including province, rural residence, and logged distance to the Somalia border; demographic characteristics, including gender, age, and education; socioeconomic measures, including rental status, monthly rent, employment, and household composition; and politically salient indicators, including religion, language, and ethnicity.

Table 5.2: Largest observed pre/post imbalances, excluding surveys from April 2

Category	Covariate	Pre mean	Post mean	Difference	SMD	Sig
Ethnicity	Kamba	0.034	0.209	0.175	0.694	***
Ethnicity	Luhya	0.177	0.036	-0.141	-0.406	***
Geography / rollout	Logged distance to Somalia	6.245	6.061	-0.184	-0.568	***
Geography / rollout	Eastern	0.083	0.231	0.147	0.466	***
Geography / rollout	North Eastern	0.010	0.075	0.065	0.419	***
Geography / rollout	Western	0.130	0.012	-0.117	-0.393	***
Geography / rollout	Rural	0.584	0.757	0.174	0.363	***
Geography / rollout	Nairobi	0.135	0.066	-0.069	-0.215	***
Other covariates	Public School	0.433	0.323	-0.111	-0.226	***
Other covariates	Private School	0.244	0.167	-0.076	-0.182	***
Other covariates	N Adults	1.719	1.585	-0.134	-0.154	**
Religion	Muslim	0.063	0.138	0.075	0.277	***

Table 5.2 reports the largest imbalances in the sample. The largest imbalances are concentrated in geography, ethnicity, rurality, religion, and logged distance to Somalia. Post-

attack respondents are more likely to be from Eastern and North Eastern provinces, more likely to be Kamba, more rural, more Muslim, and closer to the Somalia border; pre-attack respondents are more likely to be from Western, Nairobi, and Central provinces and more likely to be Luhya. Full balance diagnostics are reported in Appendix Table A.1.

5.3.2 Fieldwork Continuity Restriction: North Eastern Province

As discussed above, North Eastern Province is excluded from the main adjusted analyses because it is the location of the Garissa attack and the place where event-induced fieldwork changes are most plausible. This restriction removes the province where post-attack changes in enumerator movement, respondent availability, or fieldwork continuity would be most directly tied to the event itself.

The remaining analyses therefore focus on the broader rollout problem emphasized in this thesis: even outside the most obvious site of event-induced fieldwork disruption, preexisting survey rollout may still produce geographically structured pre/post differences.

5.3.3 Common Support

The descriptive evidence in Section 5.1 already shows that geographic coverage is uneven across the cutoff. I therefore do not treat province imbalance as a new result here. Instead, the common-support question is whether, after recognizing that geographic rollout may structure pre/post respondent composition, the data contain pre- and post-attack respondents who are comparable within substantively relevant categories likely to matter for potential outcomes.

5.3.3.1 Geographic Overlap

Because the most visible imbalances are geographic, the most direct adjustment strategy would be to compare pre- and post-attack respondents who are geographically close to one another. I therefore first assess whether location-based matching is feasible. Figure 5.3 plots the distance from each pre-attack respondent to the nearest post-attack respondent in the restricted sample. The median distance is 19 km, meaning that matching even half of pre-attack respondents to their nearest post-attack counterpart would require allowing matches at distances of roughly 20 km. This is a wide radius for assuming local comparability in potential outcomes. Geographic support is also concentrated: among the pre-attack respondents with a nearest post-attack respondent within the median distance, the nearest-neighbor matches draw on only 74 of the 381 available post-attack respondents. Thus, a location-based matching strategy would either require allowing matches over wide geographic distances or rely heavily on a small subset of post-attack locations. Even among geographically closer matches, comparability is not guaranteed: if survey teams returned to an area to complete quotas or revisit households, respondents interviewed later in the same area may still differ systematically from those interviewed earlier. Taken together, location-based matching does not provide sufficient overlap to serve as the main adjustment strategy.

The limited geographic overlap also clarifies the identifying assumption required for adjustment. If geography matters because it is related to respondent characteristics that predict potential outcomes, then adjustment must proceed through observed covariates that approximate those characteristics.

5.3.3.2 Covariate-Space Support

The adjustment strategy proceeds in two steps. First, I restrict the sample to ethnic groups with meaningful pre- and post-attack representation. I do this because ethnicity is substantively central in Kenyan politics, strongly related to regional and coalition cleavages, and visibly imbalanced across the cutoff. This restriction ensures that the main comparisons are not driven by ethnic categories observed only before or only after the attack. Second, within this restricted sample, I use covariate adjustment to address remaining observed differences across pre/post respondents. The common-support restriction therefore defines the sample

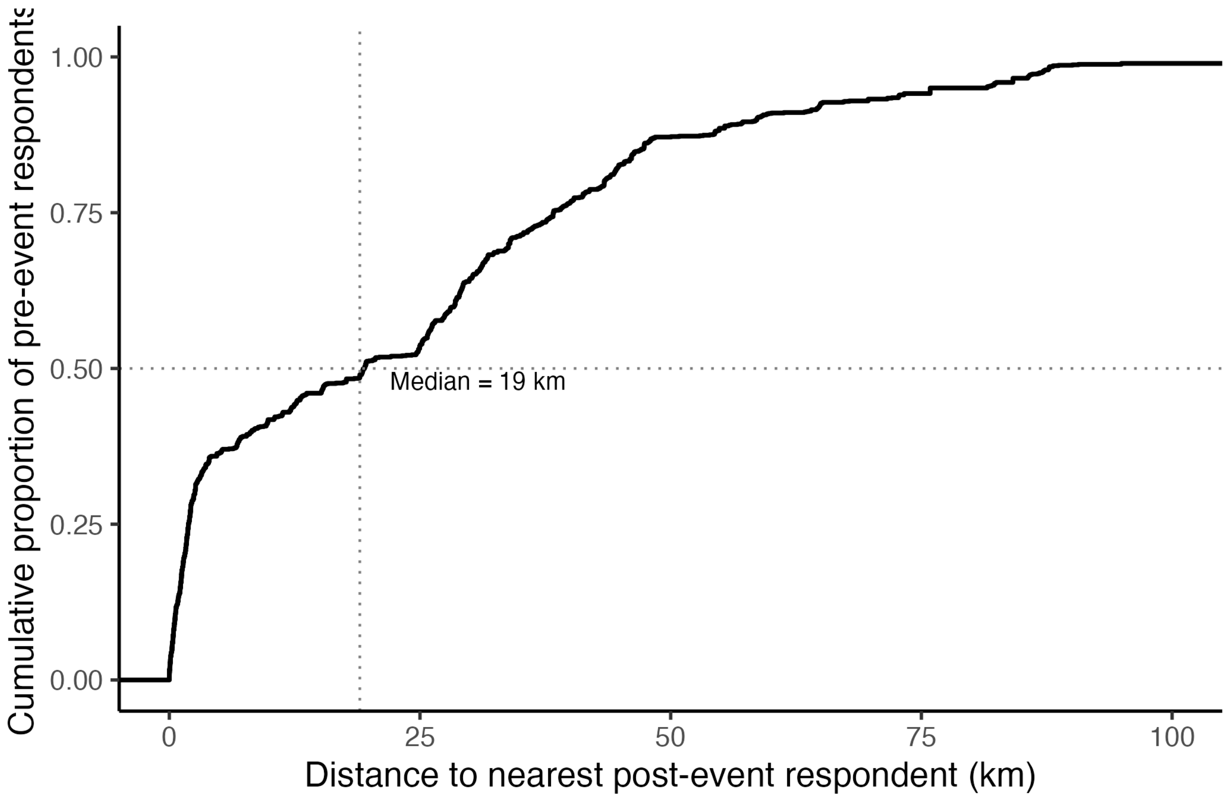

Figure 5.3: The distribution of distances from each pre-event respondent to the nearest post-event respondent.

in which comparison is possible; the adjustment model addresses observed imbalance within that sample.

This does not eliminate the possibility of unobserved geographic confounding. Instead, it shifts the design from geographic comparability to conditional comparability in covariate space. The identifying claim is that, within the observed characteristics used for adjustment, remaining pre/post status is unrelated to potential outcomes.

I retain ethnic groups with at least 10 respondents in both the pre- and post-attack periods. This rule preserves major groups with meaningful representation on both sides of the cutoff while excluding groups for which comparisons would rely on sparse data. The restriction retains Kikuyu, Kalenjin, Luhya, Luo, Kamba, Kisii, and Mijikenda respondents. Importantly, the resulting common-support sample is not a population-wide sample of all respondents in the survey; it is the subset with meaningful pre- and post-attack representation within ethnicity categories.

5.4 Outcome-Predictive Covariates and SOO-Adjusted Estimates

Having defined the common-support sample, I next estimate adjusted pre/post comparisons. These analyses rely on a conditional ignorability assumption: within the observed respondent characteristics used for adjustment, pre/post interview status is assumed to be unrelated to potential outcomes. This is a selection-on-observables claim, not a claim that the event timing alone generated as-if random assignment.

5.4.1 Covariate Selection with LASSO

The remaining question is which observed covariates to include in the adjusted models. I begin from the impact $\times$ imbalance framework developed in Chapter 3: measured characteristics are most relevant for adjustment when they are both imbalanced across pre/post status and predictive of potential outcomes. Balance checks identify whether these char-

Table 5.3: Ethnicity-based common support and inclusion decisions

Ethnicity	Pre-attack	Post-attack	Total	Common support	Decision
Kikuyu	289	79	368	TRUE	Retain
Kalenjin	208	49	257	TRUE	Retain
Luhya	241	15	256	TRUE	Retain
Luo	191	60	251	TRUE	Retain
Kamba	46	86	132	TRUE	Retain
Kisii	102	16	118	TRUE	Retain
Mijikenda	57	31	88	TRUE	Retain
Meru	46	5	51	FALSE	Exclude
Other	25	6	31	FALSE	Exclude
Turkana	23	0	23	FALSE	Exclude
Embu	20	2	22	FALSE	Exclude
Taita	12	9	21	FALSE	Exclude
Rendille	14	5	19	FALSE	Exclude
Teso	11	7	18	FALSE	Exclude
Mbeere	15	1	16	FALSE	Exclude
Maasai	10	3	13	FALSE	Exclude
Swahili	9	1	10	FALSE	Exclude
Somali	7	2	9	FALSE	Exclude
Tharaka	8	1	9	FALSE	Exclude
Arab	5	0	5	FALSE	Exclude
Kenyan only	2	3	5	FALSE	Exclude
Bajuni	1	0	1	FALSE	Exclude
Borana	1	0	1	FALSE	Exclude
Samburu	1	0	1	FALSE	Exclude
Taveta	1	0	1	FALSE	Exclude

acteristics differ across pre/post status. Outcome-prediction exercises then assess whether they are related to untreated or post-attack outcomes.

I use LASSO logistic regression as a screening tool for observed covariates that predict support for KDF operations. I estimate the model separately in the pre-attack and post-attack periods. The pre-period screen uses observed outcomes among untreated respondents as information about baseline support, $Y(0)$. The pre-period model screens for covariates associated with observed untreated outcomes, $Y_i(0) \sim X_i$, where X_i are observed pretreatment covariates. I also estimate an analogous post-attack screen, $Y_i(1) \sim X_i$, which may capture characteristics related to heterogeneous responses to the attack. I then take the union of the selected covariates as the set of outcome-relevant predictors for adjusted comparisons.

Geography and rollout-related variables, including province, rural residence, and logged distance to the Somalia border, could help diagnose how fieldwork shaped pre/post respondent composition. Other variables plausibly related to potential outcomes include demographic characteristics such as gender, age, and education; socioeconomic factors such as home rental status, monthly rent, employment type, and number of adults in the household; and politically salient indicators such as religion, language spoken at home, and ethnicity.

The summarized results of the LASSO selection are in Table 5.4. The LASSO screens use `lambda.min` from cross-validation. Coefficients are not interpreted substantively; the procedure is used only to select observed outcome-predictive covariates for adjusted specifications. The LASSO tuning parameters are reported in Appendix Table A.2.

5.4.2 SOO-Adjusted Estimate

I now estimate the SOO-adjusted pre/post association between the Garissa attack and support for KDF operations. The estimating equation is:

$$Y_i = \alpha + \tau D_i + X_i^* \beta + \varepsilon_i,$$

Table 5.4: LASSO-selected adjustment set

Screen	Selected predictors
Pre-attack outcome screen	rural, female, rent, rel_muslim, rel_catholic, public_school, self_employed, unemployed, agriculture, eth_Kalenjin, eth_Kikuyu, eth_Luhya, provinceEastern, provinceWestern
Post-attack outcome screen	female, age, rel_sda, rel_otherchristian, any_famine, public_school, employ_private, peasant, eth_Kalenjin, eth_Kikuyu, provinceCentral

where Y_i is support for KDF operations, D_i indicates that respondent i was interviewed after the attack, and X_i^* is the union of covariates selected by the pre- and post-period LASSO screens. The coefficient τ estimates the adjusted difference in KDF support between post- and pre-attack respondents in the common-support sample. This should be interpreted as a selection-on-observables estimate: it is credible only if, conditional on the included covariates, remaining pre/post status is not systematically related to potential outcomes.

The results are reported in Table 5.5. The SOO-adjusted estimate is negative: respondents interviewed after the Garissa attack were about 7 percentage points less likely to support KDF operations than otherwise comparable respondents interviewed before the attack. This estimate is only marginally distinguishable from zero ($p = 0.069$) and is not statistically significant at the 5% level. Substantively, the direction of the estimate is consistent with a decline in KDF support after the attack, but the evidence is not strong enough to treat this adjusted association as precise or definitive.

Table 5.5: SOO-adjusted regression results

Covariate	Estimate	SE	p-value	Sig
Intercept	0.525	0.076	0.000	***
Post-attack	-0.070	0.038	0.069	†
Rural	0.067	0.036	0.065	†
Female	-0.117	0.031	0.000	***
Rents home	-0.039	0.040	0.326	
Muslim	-0.206	0.080	0.010	*
Catholic	-0.061	0.035	0.083	†
Public school	-0.028	0.033	0.402	
Self-employed	0.066	0.041	0.106	
Unemployed	-0.038	0.052	0.466	
Agriculture	0.029	0.044	0.508	
Kalenjin	0.100	0.044	0.024	*
Kikuyu	0.154	0.048	0.001	**
Luhya	-0.037	0.055	0.497	
Eastern province	-0.054	0.069	0.435	
Western province	-0.179	0.068	0.009	**
Age	0.000	0.001	0.840	
Seventh Day Adventist	-0.022	0.052	0.677	
Other Christian	0.057	0.051	0.263	
Famine exposure	-0.028	0.032	0.372	
Private-sector employed	-0.042	0.050	0.402	
Peasant/farmer	0.045	0.072	0.532	
Central province	0.018	0.054	0.745	
Num. obs.	1078			
R-squared	0.100			

5.4.3 Heterogeneity by Coalition

The pooled SOO-adjusted estimate asks whether average KDF support changed after the Garissa attack within the supported sample. But the central concern of this chapter is not only whether the pooled estimate changes after adjustment, but whether the pre/post comparison combines respondents whose political relationship to the KDF mission differs in substantively important ways. Geographic rollout matters partly because places differ in their political composition. In the Kenyan case, one especially important dimension of that composition is ethnic-coalition alignment.

The common-support restriction above retains ethnic groups with meaningful representation on both sides of the cutoff. I now aggregate those supported ethnic groups into government- and opposition-aligned coalitions. This aggregation is theoretically motivated by the political context discussed in Chapter 4: the KDF mission in Somalia was associated with the Kenyatta-Ruto government's counterterrorism strategy, and respondents aligned with the governing coalition may have interpreted the Garissa attack differently from respondents aligned with the opposition.

I code coalition alignment using majority 2013 vote patterns among the ethnic groups retained by the common-support restriction. Kikuyu and Kalenjin respondents are coded as government-aligned, while Luo, Luhya, Kamba, Kisii, and Mijikenda respondents are coded as opposition-aligned. Coalition alignment is included directly in the interaction model and I exclude ethnic indicators from the adjustment covariates.

$$Y_i = \alpha + \tau D_i + \gamma Opposition_i + \delta(D_i \times Opposition_i) + X_i^* \beta + \varepsilon_i,$$

where $Opposition_i$ indicates that respondent i belongs to an opposition-aligned ethnic group, and X_i^* includes the selected non-ethnicity adjustment covariates. The coefficient δ captures whether the post-attack shift differs between opposition- and government-aligned respondents.

Table 5.6: Coalition-specific adjusted post-attack estimates

Group / contrast	Estimate	SE	p-value	Sig
Government coalition	0.030	0.053	0.580	
Opposition coalition	-0.114	0.045	0.013	*
Difference: Opposition - Government	-0.143	0.069	0.039	*

The coalition-level estimates show that the pooled decline is concentrated among opposition-aligned respondents. Among government-coalition respondents, post-attack interview status is associated with a small and statistically insignificant 3.0 percentage-point increase in KDF support. Among opposition-coalition respondents, post-attack interview status is associated with an 11.4 percentage-point decline in KDF support. The difference between these shifts is 14.3 percentage points and statistically significant at the 5% level. This pattern is consistent with the interpretation that the Garissa attack did not produce a uniform response across politically salient groups: support remained essentially unchanged among government-aligned respondents, while opposition-aligned respondents became less supportive of the KDF mission.

5.5 Placebo Cutoff Analysis

I next compare the main SOO-adjusted estimate to two placebo cutoffs in time. The goal is to assess whether similar adjusted pre/post differences appear at dates when no focal event occurred. I estimate the same adjusted specification within the pre-attack period using a false cutoff between March 29 and March 30, and within the post-attack period using a false cutoff between April 3 and April 4. These placebo tests are not definitive, especially because the within-period samples are smaller, but they provide a diagnostic for whether remaining fieldwork composition shifts can generate apparent pre/post differences away from the actual event date.

Table 5.7: Placebo cutoff analysis results

Test	Estimate	SE	p-value	N
Main estimate: April 2 cutoff	-0.070	0.038	0.069	1078
Pre-period placebo: March 29/30 split	0.097	0.038	0.010	810
Post-period placebo: April 3/4 split	0.004	0.072	0.961	268

The placebo results underscore the central methodological point of this chapter. The main adjusted estimate suggests a 7 percentage-point decline in KDF support after the Garissa cutoff, but the pre-period placebo estimate is positive, statistically significant, and larger in absolute value than the main estimate. Because no focal event occurs at this placebo cutoff, this result suggests that residual fieldwork-related composition shifts remain consequential even after the common-support restriction and covariate adjustment. The post-period placebo estimate is close to zero, so the evidence is not uniformly pessimistic. Still, the pre-period placebo cautions against interpreting the pooled SOO-adjusted estimate as a clean causal effect.

5.6 Discussion

All estimates in this chapter should be read as intent-to-treat estimates of post-attack interview status. The data do not measure whether respondents were aware of the attack, how much information they had, or how salient the event was at the time of interview.

The common-support restriction and SOO adjustment improve the comparability of pre- and post-attack respondents, but they do not eliminate the core identification concern. The location-based matching check shows that geographic overlap is limited, and the placebo cutoff analysis indicates that fieldwork composition can still generate sizable adjusted differences away from the actual event date. The pre-period placebo estimate is especially important: because no focal event occurred at that cutoff, the result suggests that residual fieldwork rollout remains capable of producing apparent pre/post differences even after adjustment. For this reason, the pooled SOO-adjusted estimate should be interpreted as suggestive rather than definitive causal evidence.

The coalition results add a substantively important pattern. The adjusted decline in KDF support is concentrated among opposition-aligned respondents, while government-aligned respondents show no significant change. This is consistent with the idea that the Garissa attack may have been interpreted differently across politically salient groups. However, the placebo results discipline this interpretation. Without the placebo analysis, the coalition pattern might appear to provide stronger causal evidence than the design can support. The pre-period placebo shows that similarly meaningful-looking adjusted differences can arise at non-event cutoffs, suggesting that residual rollout-related composition remains a serious concern.

The analytic restrictions also narrow the population to which the estimates apply. Dropping April 2 removes observations with ambiguous treatment timing. Dropping North Eastern Province removes the province where fieldwork continuity is least defensible. Restricting to ethnic groups with meaningful pre- and post-attack representation further limits the sample to respondents for whom covariate-space comparisons are supported by the data. These restrictions may improve internal validity, but the resulting estimates should not be interpreted as population-average effects for all Kenyan respondents in the original survey. They apply most directly to the common-support sample defined by the fieldwork and covariate-support restrictions.

Taken together, the Garissa analyses suggest a possible decline in KDF support after the attack, concentrated among opposition-aligned respondents. But the evidence is not strong enough to support a definitive causal interpretation. I therefore interpret the results as suggestive evidence of attitude change within the restricted common-support sample, and as stronger evidence for the methodological claim of this thesis: even when adjusted estimates are substantively coherent, fieldwork rollout can still shape the comparison in ways that limit causal inference.

CHAPTER 6

Conclusion

This thesis has argued that unexpected events during survey fieldwork can provide useful causal leverage, but only under more limited conditions than many applications imply. The key point is that the event date does not by itself determine whether pre- and post-event respondents are comparable. Respondents are observed before or after the event through the interaction between the event date and the realized fieldwork sequence. When fieldwork rollout is spatially or demographically structured, the event cutoff may divide not only time, but also geography and respondent composition in ways related to potential outcomes.

This does not mean that UESD designs should be rejected. Rather, they should be understood as conditional designs: often closer to selection-on-observables than to randomized assignment, but with additional diagnostic leverage. In this sense, UESD is "SOO+." Like ordinary selection-on-observables designs, it requires researchers to defend the claim that pre/post status is independent of potential outcomes conditional on observed covariates. But unlike many observational designs, the fieldwork sequence associated with pre/post status is often more visible. Researchers can observe survey dates, fieldwork locations, geographic movement, balance patterns, common support, and placebo cutoffs in time. These features do not prove ignorability, but they make the credibility of the design more diagnosable.

The Garissa attack illustrates both the promise and limits of this approach. The attack occurred during an active survey, creating an attractive pre/post design. Yet the survey rollout produced substantial geographic and compositional differences across the cutoff. In this setting, those compositional differences were politically meaningful: the adjusted decline in KDF support was concentrated among opposition-aligned respondents, while government-aligned respondents showed little change. A regression discontinuity interpretation was not credible because adjacent survey dates did not necessarily contain comparable respondents. Location-based matching was also limited because geographic overlap was thin. Restricting the sample to common support and adjusting for outcome-predictive covariates improved the comparison, but placebo cutoffs showed that residual fieldwork composition could still generate sizable apparent effects away from the actual event date.

The broader lesson is that design checks should not be treated as a checklist applied after a UESD design is assumed valid. Maps, balance tables, common-support checks, outcome-prediction exercises, and placebo cutoffs help determine how the comparison should be interpreted: as a plausibly as-if random pre/post comparison, as a conditional selection-on-observables comparison, as an estimate restricted to a supported subset, or as a comparison too affected by residual confounding or limited overlap to sustain causal claims. Once researchers restrict to common support or condition on covariates, the estimate should no longer be described as a simple population-wide effect of the event. Often, it is an adjusted intent-to-treat estimate for the subset of respondents whose observed covariate profiles support a credible pre/post comparison.

This concern is especially acute in face-to-face surveys with spatially sequenced fieldwork. When survey teams move through regions, municipalities, neighborhoods, or clusters over time, the event cutoff may separate respondents not only by interview date but also by place. The problem is not observed geographic imbalance itself. The deeper concern is that geographic rollout can produce pre- and post-event groups that differ in what their outcomes would have been absent the event. Geography can proxy for ethnicity, political alignment, security exposure, socioeconomic conditions, media environments, and other determinants of outcomes. Researchers should therefore ask not only whether observed geography is balanced, but whether the rollout-related confounding could create a plausible backdoor path between interview timing and potential outcomes.

UESD approaches remain useful because they exploit real-world shocks that could not be randomized and often occur in substantively important settings. But they are not as powerful as the language of natural experiments sometimes suggests. Their strength lies less in automatic as-if random timing and more in the transparency of the fieldwork process. Used carefully, UESD can provide credible and policy-relevant evidence while making clear the scope and limits of the comparison. Used casually, it risks turning survey logistics into causal effects. The task for researchers is therefore not to abandon UESD, but to make realized fieldwork central to identification, scope claims to the supported sample, interpret estimates as ITT unless exposure is measured, and treat design checks as evidence about the credibility of the comparison rather than optional robustness checks.

APPENDIX A

Additional Tables and Figures

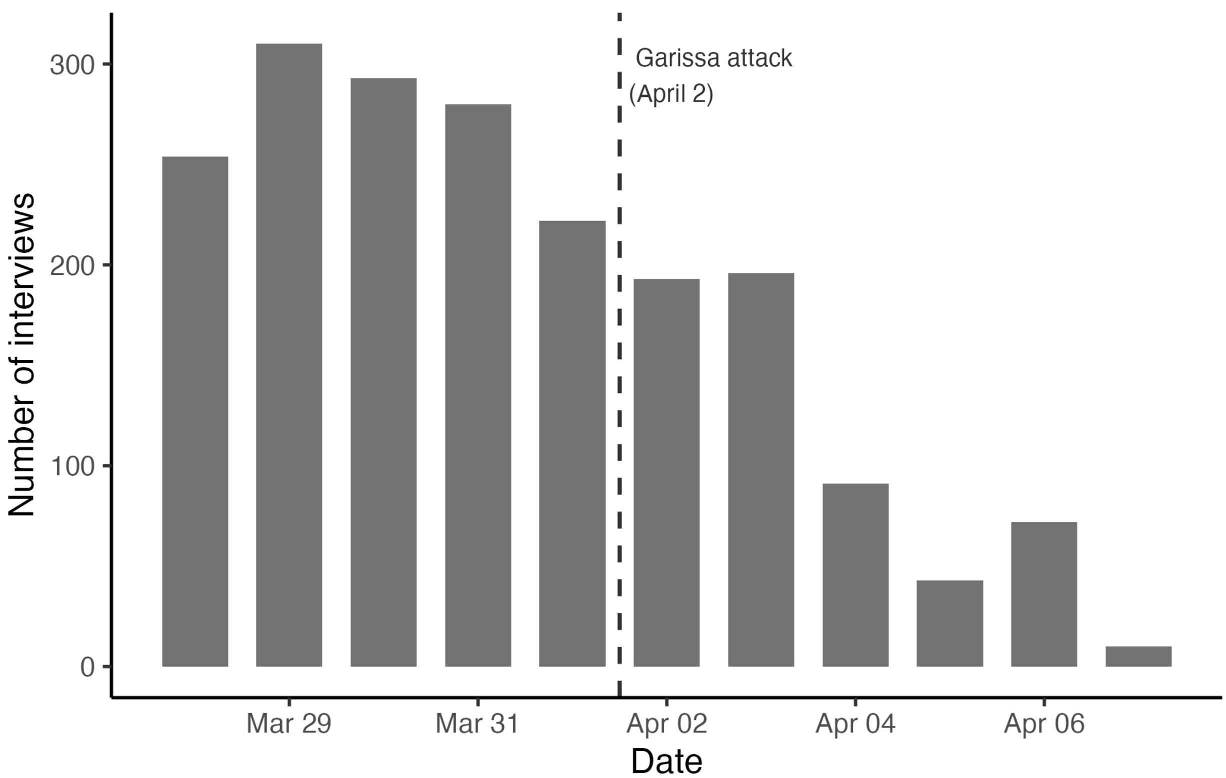

Figure A.1: Survey responses by day.

Table A.1: Full-sample pre/post balance, excluding April 2

Category	Covariate	Pre mean	Post mean	Difference	SMD	Sig
Geography / rollout	Central province	0.149	0.097	-0.052	-0.152	**
Geography / rollout	Coast province	0.093	0.124	0.031	0.104	†
Geography / rollout	Eastern province	0.083	0.231	0.147	0.466	***
Geography / rollout	Nairobi province	0.135	0.066	-0.069	-0.215	***
Geography / rollout	North Eastern province	0.010	0.075	0.065	0.419	***
Geography / rollout	Nyanza province	0.135	0.155	0.020	0.057	
Geography / rollout	Rift Valley province	0.265	0.240	-0.025	-0.056	
Geography / rollout	Western province	0.130	0.012	-0.117	-0.393	***
Geography / rollout	Rural	0.584	0.757	0.174	0.363	***
Geography / rollout	Logged distance to Somalia	6.245	6.061	-0.184	-0.568	***
Ethnicity	Kalenjin	0.153	0.119	-0.034	-0.097	†
Ethnicity	Kamba	0.034	0.209	0.175	0.694	***
Ethnicity	Kikuyu	0.213	0.192	-0.021	-0.052	
Ethnicity	Luhya	0.177	0.036	-0.141	-0.406	***
Ethnicity	Luo	0.141	0.146	0.005	0.015	
Ethnicity	Ethnicity missing/other	0.283	0.299	0.016	0.035	
Religion	Muslim	0.063	0.138	0.075	0.277	***
Religion	Catholic	0.280	0.250	-0.030	-0.066	
Religion	Protestant	0.304	0.296	-0.008	-0.017	
Religion	Evangelical	0.132	0.097	-0.035	-0.107	*
Religion	Seventh Day Adventist	0.088	0.095	0.007	0.025	
Religion	Other Christian	0.097	0.080	-0.017	-0.059	
Language	Swahili only	0.322	0.286	-0.035	-0.076	
Language	English and Swahili	0.663	0.655	-0.008	-0.016	
Demographics	Female	0.470	0.398	-0.072	-0.145	**
Demographics	Age	32.644	31.934	-0.709	-0.061	
Demographics	Education	9.908	10.454	0.546	0.142	*
Socioeconomic / household	Rents home	0.439	0.459	0.019	0.039	
Socioeconomic / household	Monthly rent	1673.142	1580.097	-93.045	-0.031	
Socioeconomic / household	Total income	15360.559	14199.029	-1161.530	-0.073	
Socioeconomic / household	Number of adults	1.719	1.585	-0.134	-0.154	**
Socioeconomic / household	Registered to vote	0.783	0.840	0.057	0.141	**
Employment / livelihood	Self-employed	0.426	0.400	-0.026	-0.052	
Employment / livelihood	Unemployed	0.177	0.160	-0.017	-0.045	
Employment / livelihood	Private-sector employed	0.121	0.165	0.044	0.132	*
Employment / livelihood	Peasant/farmer	0.092	0.070	-0.022	-0.077	
Employment / livelihood	Agriculture	0.227	0.180	-0.048	-0.116	*
Schooling / exposure	Public school	0.433	0.323	-0.111	-0.226	***
Schooling / exposure	Private school	0.244	0.167	-0.076	-0.182	***
Schooling / exposure	Famine exposure	0.559	0.556	-0.003	-0.007	

Table A.2: LASSO tuning parameters and selected covariate counts

Screen	N	lambda.min	lambda.1se	Number selected
Pre-attack outcome screen: Y(0) $\sim$ X	810	0.019	0.036	14
Post-attack outcome screen: Y(1) $\sim$ X	268	0.039	0.062	11

REFERENCES

David M. Anderson and Jacob McKnight. Kenya at war: Al-Shabaab and its enemies in Eastern Africa. *African Affairs*, 114(454):1–27, 2015.

Macarena Ares and Enrique Hernández. The corrosive effect of corruption on trust in politicians: Evidence from a natural experiment. *Research & Politics*, 4(2):1–8, 2017.

Andrews Atta-Asamoah and Emmanuel Kisiangani. Implications of Kenya's military offensive against Al-Shabaab. ISS Today, Institute for Security Studies, 2011.

Andrew Bertoli, Laura Jakli, and Henry Pascoe. Analyzing the impact of events through surveys: Formalizing biases and introducing the dual randomized survey design. *Political Science Research and Methods*, 14(2):255–275, 2026.

Travis B. Curtice. How repression affects public perceptions of police: Evidence from a natural experiment in Uganda. *Journal of Conflict Resolution*, 65(10):1680–1708, 2021.

Peter Thisted Dinesen and Mads Meier Jæger. The effect of terror on institutional trust: New evidence from the 3/11 Madrid terrorist attack. *Political Psychology*, 34(6):917–926, 2013.

Karen E. Ferree, Clark C. Gibson, and James D. Long. Voting behavior and electoral irregularities in kenya's 2013 election. *Journal of Eastern African Studies*, 8(1):153–172, 2014. doi: 10.1080/17531055.2013.874157.

Timothy Frye and Ekaterina Borisova. Elections, protest, and trust in government: A natural experiment from Russia. *The Journal of Politics*, 81(3):820–832, 2019.

Guido W. Imbens and Donald B. Rubin. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction.* Cambridge University Press, 2015.

Luckystar Miyandazi. Kenya's military intervention in Somalia: An intricate process. ACCORD Policy & Practice Brief 19, African Centre for the Constructive Resolution of Disputes, 2012.

Jordi Muñoz, Albert Falcó-Gimeno, and Enrique Hernández. Unexpected event during survey design: Promise and pitfalls for causal inference. *Political Analysis*, 28(2):186–206, 2020.

Oscar Gakuo Mwangi. State collapse, Al-Shabaab, islamism, and legitimacy in Somalia. *Politics, Religion & Ideology*, 13(4):513–527, 2012.

Christof Nägel and Amy E. Nivette. Unexpected events during survey design and trust in the police: A systematic review. *Journal of Experimental Criminology*, 19:891–917, 2023.

Paul R. Rosenbaum and Donald B. Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 1983.

Marc Sangnier and Yanos Zylberberg. Protests and trust in the state: Evidence from African countries. *Journal of Public Economics*, 152:55–67, 2017.

Chris G. Sibley, Lara M. Greaves, Nicole Satherley, Marc S. Wilson, Nickola C. Overall, Carole H. J. Lee, Petar Milojev, Joseph Bulbulia, Danny Osborne, Taciano L. Milfont, Carla A. Houkamau, Isabelle M. Duck, Rachel Vickers-Jones, and Fiona K. Barlow. Effects of the COVID-19 pandemic and nationwide lockdown on trust, attitudes toward government, and well-being. *American Psychologist*, 75(5):618–630, 2020.

Anne Speckhard and Ardian Shajkovci. The Jihad in Kenya: Understanding Al-Shabaab recruitment and terrorist activity inside Kenya—in their own words. *African Security*, 12 (1):3–61, 2019.

Clair White, David Weisburd, and Sarah Wire. Examining the impact of the Freddie Gray unrest on perceptions of the police. *Criminology & Public Policy*, 17(4):829–858, 2018.