Deterrence Through Response Curves:

An Empirical Analysis of the Gaza-Israel Conflict

Eli Berman[†], Prabin B. Khadka[‡], Danny Klinenberg[§], and Esteban F. Klor[¶]

April 2024

We use response curves in a repeated game to formalize key aspects of integrated deterrence: escalation, de-escalation, incomplete deterrence, and deterrence by denial. In this approach, episodes of violence are due to the interaction of response curves, which follow a strategic logic of disincentivizing opponents from attacking, through both deterrence and compellence. To maintain credibility in future episodes both sides punish attacks, disincentivizing larger attacks and yielding nonviolent lulls. We empirically estimate those curves using detailed incident data from the Israel-Gaza conflict between 2007 and 2017. Our estimates match the dynamics of the raw data: very frequent episodes of low lethality violent exchange. Response curves are stable and exhibit a posture consistent with incomplete deterrence: i.e., episodes de-escalate, but not to complete nonviolence in equilibrium. Major Israeli military operations shift the Gazan response curve inwards, yielding a less violent equilibrium. Iron dome missile defense does not.

Keywords: Conflict, Deterrence; Repeated games, Israel – Gaza conflict;

[†] Department of Economics, University of California at San Diego; elib@ucsd.edu; [‡] Department of Government, University of Essex; prabin.khadka@essex.ac.uk; [§] Institute of Global Conflict and Cooperation, and University of California San Diego; dklinenberg@ucsb.edu; [¶] Department of Economics, The Hebrew University of Jerusalem and The Institute for National Security Studies (INSS); eklor@huji.ac.il;

We thank Anastasiia Faikina, David Rigg, and Gal Zweigel for superb research assistance. We thank participants at seminars at The Hebrew University, the University of Calgary, West Point and Carleton University for valuable discussions and suggestions. We also thank participants at numerous conferences for helpful comments and feedback. We acknowledge the support of the Minerva Research Institute (FA9550-23-1-0437). Esteban F. Klor thanks the Israel Science Foundation (grant No. 1061/18) for financial support.

1. Introduction

Deterring adversaries is fundamental to national security. Yet deterrence is often incomplete: even the most powerful actors suffer a plethora of low-level attacks and provocations, from espionage, cyber-attacks, and election interference to proxy terrorism and missile tests. We use the analytical device of response curves in repeated games to examine key aspects of deterrence relevant to modern conflict: escalation, de-escalation, incomplete deterrence, and deterrence by denial. Response curves, (equivalently, reaction curves) draw on a rich literature in noncooperative game theory (Schelling 1960 and 1966; Fudenberg and Tirole 1991).[1]

Our approach allows us to empirically estimate response curves and to test theoretical predictions about deterrence by denial, using high-frequency data from the Israel-Gaza conflict, covering the period 2007-2018. That period was characterized by long periods of relatively low-intensity violence at high frequency (about an episode of violent exchange every ten days) interrupted by three large Israeli incursions into Gaza, aimed at destroying Hamas' capacity to attack.

In our model, episodes of violence are due to the interaction of stable response curves. Response curves follow a strategic logic of disincentivizing opponents from attacking, through both deterrence and compellence. To maintain credibility both sides punish attacks, which result in periods of relative quiet and disincentivize larger attacks. Deterrence can be incomplete (i.e., violence occurs in equilibrium) when at least one adversary faces high net costs of suppressing low levels of violence, while another has positive net returns to low-level violence. We don't take a strong stand on the underlying social welfare functions from which response curves might be derived. A response curve approach to analyzing episodes of conflict can encompass both rationalist and other explanations for conflict, such as leaders appeasing domestic audiences who might seek retribution (Fearon 1994), or conflict due to psychological biases or over-responses of combatants and decision-makers (Jervis, Lebow and Stein 1985).[2]

---

[1] The idea of partial deterrence is developed in Freedman (2004).

[2] Building on the original insights of Schelling (1960, 1966), Kahn (1960), Mearsheimer (1983) and Powell (1990), several formal game-theoretic models of deterrence have been proposed to tie up loose ends regarding complexities of credibility. Subsequent studies have also yielded additional significant theoretical breakthroughs, providing rigorous game-theoretic foundations for the analysis of current policy questions such as how to promote state building, or whether and how to militarily intervene to induce regime change (Fearon 1997 and 2002; Powell 1999, 2003, 2012; Myerson 2008; Baliga and Sjöström 2009 and Chassang and Padró i Miquel 2010).

Much game-theoretic work focuses on a single crisis where the threat of conflict or escalation is greatly influenced by some unobserved characteristic of an opponent, which shapes their response curve through perceived costs and benefits of attack. The unobserved characteristic could be the resolve of the opponent (i.e., the importance that the opponent gives to a particular issue under bargaining) or her military ability, which would in turn affect her (unknown) propensity to escalate attacks. This approach seems less compelling in long-standing conflicts such as Israel-Gaza, where opponents have plenty of opportunity to learn each-other's characteristics.

The period of incomplete deterrence that we study came to a stunning end on October 7, 2023, so some context is necessary before turning to empirical results. First, as the literature on strategic surprise points out, deterrence is achieved when an adversary perceives that costs of a violent attack exceed benefits. Yet that perception, in the mind of an adversary, might well be hidden. Specifically, an adversary can achieve surprise (once deterrence has failed), by masking their true perception of costs and benefits until they attack (e.g., Chassang and Padro i Miquel, 2008). That seems to have been what Hamas leadership in Gaza achieved during the two years before the October 7th attacks.[3]

Our data describe a period including Yahya Sinwar's election to lead Hamas in early 2017, but before documented preparations for the October 7, 2023 terrorist attack, which apparently began sometime in 2021 (though a similar plan existed already in 2014).[4] During our period of analysis (2007-2018) our working assumption is that Hamas' response curves were not a façade, but instead reflected a stable underlying welfare function.

Those data are drawn from daily reports as recorded by the United Nations Field Security Office. Those reports describe security incidents in one or two sentences of text, followed by a

---

[3] We lack corroborated evidence on why Hamas' leadership in Gaza (principally Yahya Sinwar and Mohammed Deif) perceived that the costs of a large attack on civilians no longer exceeded benefits by October 2023. Speculation based on comments by Hamas' leadership abroad include: (i) increased fear of a peace agreement between Israel and Saudi Arabia that would favor their political rivals in the Palestinian Authority, and doom Hamas' stated goal of liberating Palestine; and (ii) a calculation that a very predictable toll in civilian casualties and casualties among their own fighters, massive as it might be, would be justified by progress in the overarching objective of liberating the holy places (Hubbard, Ben and Maria Abi-Habib, 2023).They may have also believed that Hamas' persistent accumulation of an infrastructure of tunnels, fortifications and mobilized fighters would deter the IDF from a massive military incursion, or protect them from such an attack.

[4] Bergman and Goldman, 2023; Sivan Hila'I, 2023.

description of any damage or casualties. Each description is preceded by a time stamp accurate to 5 minutes, and is organized into four geographical areas. We code those data into “actions” (multiple attacks using air strikes, small arms fire, shelling and incursions by the IDF, and using small arms fire, rocket fire and mortar fire by Gazan militants). We organize those actions sequentially into episodes of attack and counterattacks, separated by lulls. (Coding procedures are informed by social media of perpetrators.) We calculate predicted damage suffered by each side from a given attack, by associating each type of attack with actual casualties.

That approach yields a very high number of relatively low casualty episodes: 1,266 over our eleven years of data. The mean episode lasts one or two days, but is followed by only eight days of lull. Episodes average 0.9 Gazan fatalities, and 0.02 Israeli fatalities.

Using those data we estimate response curves, which map expected damage suffered into expected damage caused. Both sides’ response is consistent with mixed strategies, in the sense that they do not always respond to an attack, but their probability of response increases in expected damage suffered. Expected response curves are upwards sloping (indicating a deterrent posture), i.e., expected damage caused increases in damage suffered. Importantly, they are concave, so that they yield a single stable equilibrium. And their slopes are shallow enough that episodes de-escalate from high violence attacks back to equilibrium. Unfortunately, that equilibrium is at positive damage for both sides, a consequence of both intercepts being far from the origin. That is, our estimates exhibit incomplete deterrence in equilibrium. In particular, zero violence is not an equilibrium.

Major military operations by Israel shift the Gazan response curve to lower levels of violence, illustrating deterrence by denial, and yielding a less violent equilibrium. The introduction of the Iron Dome missile defense system does not.

Studies of conflict dynamics (see, for example, Powell 2012 or Leventoğlu and Slantchev 2007) are typically informed by a collection of cases. Our ability to estimate response curves using repeated interactions between the same opponents provides a unique empirical and analytical opportunity to illustrate both dynamics and equilibria in a stable system.

In Economics the natural analogy to our estimation approach would be a sequence of pricing decisions by duopolists in a repeated game (e.g., Bajari et al, 2007). In Political Science or Sociology this approach follows on the "sequential analysis" tradition (Lichbach 1987, Heise, 1989). When applied to two-sided conflict this approach tends to use a tit-for-tat repeated game as a model (Axelrod, 1984) and codes observational data to fit the structure of alternating sequential play in a repeated game. Moore (1998) provides an example from the literature on repression of violent protests. Our innovations –which are enabled by data quality, are in modeling both probability of attack and attack intensity, and in estimating equilibria.

One major contribution is to test two major assumptions of the repeated game approach, that response functions are stable, both within and between episodes. That test requires some nontrivial econometric analysis of serial correlation in sequences of actions and episodes, which we explain below. While those assumptions are technically rejected, we find that response functions are stable enough to be an excellent analytical tool in studying mutual deterrence.

These results differ from those in the empirical literature on the Israel-Palestine conflict. Where we find de-escalation within a few days on average, that literature finds responses to an initial provocation that last over a month. Seminal empirical studies of the Israel-Palestine conflict have studied Israeli responses to Palestinian attacks (Jaeger and Paserman 2006, 2008) and Palestinian responses to Israeli attacks (Haushofer et al, 2010). Yet these efforts lacked the necessary frequency of measurement to estimate response curves, solve equilibria, or fully describe dynamics, relying instead on reduced form approaches, principally Vector Autoregression (VAR). Haushofer et al. (2010) extended the Jaeger and Paserman (2008) analysis to include non-lethal acts of retaliation (e.g., Qassam rocket fire) using data from Gaza as well as from the West Bank. They find that optimal lag length varies for each outcome. They conclude that Israeli military actions against Palestinians lead to escalation rather than incapacitation, and that Palestinians are in fact reacting to Israeli behavior. Golan and Rosenblatt (2011) comment on Haushofer, Biletzki, and Kanwisher (2010) showing that the empirical response function of Jaeger and Paserman (2008) can be misleading, and demonstrating sensitivity of results to choosing lag lengths jointly or individually. Finally, Asali, Abu-Qarn, and Beenstock (2017) revisit Jaeger and Paserman (2008) focusing on modeling the problem nonlinearly. They show sensitivity of the Jaeger and Paserman (2008) results to modeling choices, and that VAR innovations are not normally distributed, leading them to estimate nonlinear vector autoregressions. They conclude that both sides were reacting to the other's violence, as did Haushofer et al (2010). We expand on that conclusion by describing full response curves which capture equilibrium strategies, and solve for equilibrium levels of violence.

In a technical extension we find, relative to that literature, that a VAR approach suffers from bias due to temporal disaggregation, which can become extreme if response lags are much longer or shorter than frequency of measurement (in "calendar" time). We illustrate through simulation that VAR may overestimate duration of response, potentially allowing false inference about impulses causing escalation (Klinenberg et al, 2024).

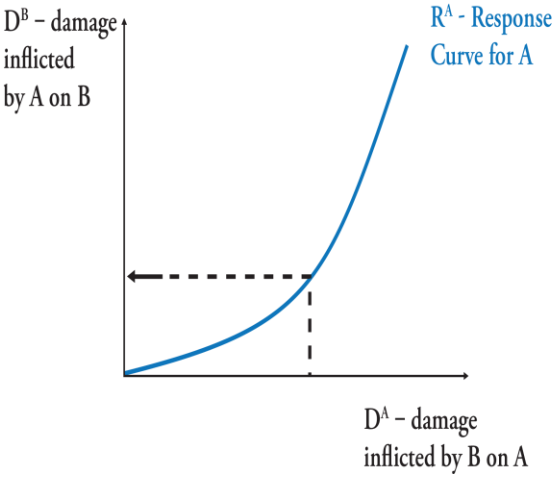
We proceed from deterministic to stochastic response curves in Section Two, explain our data in Section Three, and how we code it in Section Four. Our approach to creating a predicted damage measure out of casualties is covered in Section Five, a necessary prelude to estimating response curves in Section Six. Section Seven explores the effects of major operations in Gaza, as well as that of the Iron Dome missile defense system, and Section Eight concludes.

2. Response Curves and Deterrence

We begin by analyzing escalation and de-escalation to equilibrium in a two-sided repeated game within a deterministic model, then generalize to the stochastic case in which opponents adopt mixed strategies. Putting aside for a moment the welfare calculations that underly these responses, we start by defining terms and characterizing equilibria.

Two sides A and B suffer damage inflicted by the other, D^A and D^B, respectively. A's response is illustrated in Figure 1: A responds to damage suffered D^A (measured on the horizontal axis) with $D^B = R^A(D^A)$ (on the vertical).

Figure 1: Response Curve

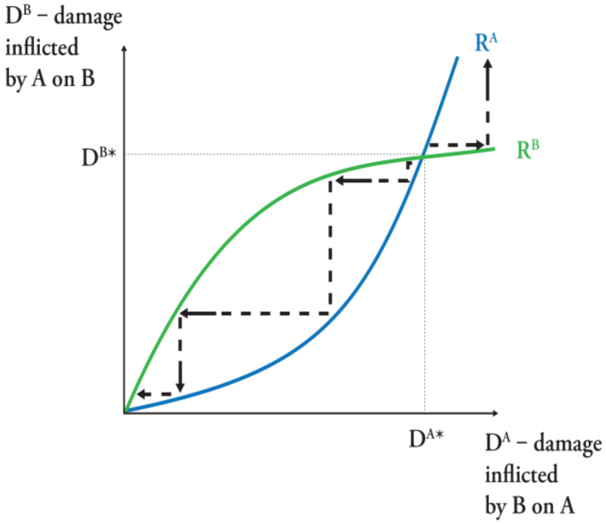


B responds with $D^A = R^B(D^B)$, as illustrated by the additional green line in Figure 2 (bearing in mind that B suffers damage measured on the vertical axis, and responds with damage measured on the horizontal axis). In the technical terms of modern game theory, they are engaged in a sequential, noncooperative, repeated, negative-sum, full-information game, with full commitment.[5]

---

[5] Ruling out simultaneous play and incomplete information simplifies our analysis. For precise implications of the full set of assumptions see Dixit and Skeath (1999, Chapter 2).

Figure 2: De-Escalation to Equilibrium



In Figure 2, both A and B have deterrent postures, by which we mean that their response increases in damage incurred, i.e., imposing increasing costs in response to increased damage. We assume that each side derives response curves strategically to maximize a hypothesized welfare function subject to technological and political constraints. While we observe damage, we don't directly observe the underlying disutility associated with it in each side's welfare function. Nor do we observe the costs each side incurs in inflicting damage on the other, or their other constraints, treaty obligations or other considerations involving third parties (allies or enemies) –all of which could influence the optimal response curve. Analytically, this is analogous to observing supply and demand curves without observing production or utility functions, with the added consideration that a response curve is a strategic choice which takes into account the opponent's response curve. We will make assumptions below on the welfare function as necessary to generate predictions.[6]

2A. Escalation and de-escalation

A sequence of responses and counter-responses is de-escalating if $R^B(R^A(D^A)) < D^A$ and $R^A(R^B(D^B)) < D^B$, at some point (D^A , D^B). I.e., a sequence of attacks that decline in damage. Those two conditions are equivalent (assuming differentiable response functions), since both require that

---

[6] We do assume that each side is a unitary actor, an assumption relaxed in Nanes (2019) discussing the same conflict, and in Dixit and Skeath's (1999) treatment of the Cuban missile crisis, pp. 446-450.

the product of derivatives is less than one [$R^{B'}(\cdot) \times R^{A'}(\cdot) < 1$]. A sequence of attacks escalates if the opposite occurs, i.e., $R^{B'}(\cdot) \times R^{A'}(\cdot) > 1$.

The serrated lines in the southwest corner of Figure 2 illustrate a de-escalating sequence and an escalating sequence. Where does the de-escalating sequence end? At a fixed point (D^{Af}, D^{Bf}), $R^B(R^A(D^{Af})) = D^{Af}$, $R^A(R^B(D^{Bf})) = D^{Bf}$, we have an equilibrium, such as the point (0,0) in Figure 2, where the two response curves intersect. That equilibrium is stable if the sequence de-escalates for D^A just above D^{Af} (and escalates just below). In Figure 2, (0,0) is a *stable* fixed point because the slopes of the two response curves imply de-escalation for levels of damage just above zero. (More precisely, the green line is steeper than the blue, since $R^{B'}(\cdot) \times R^{A'}(\cdot) < 1$, so $R^{A'}(\cdot) < 1 / R^{B'}(\cdot)$).

In contrast (D^{A*}, D^{B*}), the crossing point of the two response curves in the northeast (where the 'tail of the fish' begins), is an *unstable* equilibrium, since the slope of the blue line exceeds that of the green. For example, an attack on A with damage exceeding D^{A*} would invite a responding attack on B that exceeds D^{B*}, and an escalating sequence of attacks would follow, as illustrated by the serrated lines northeast of the intersection point.

Stable response curves in repeated episodes (i.e., sequences of attacks and responding counterattacks) would imply a dynamic equilibrium of repeated, de-escalating episodes in positive intervals bounded by D^{A*} and D^{B*}. Damage incurred at or exceeding D^{A*} or D^{B*}, respectively, would imply escalation. In this sense escalation is an implication of the shape of both response curves; in Figure 2 an attack escalates the system by inducing a response northeast of the unstable equilibrium (D^{A*}, D^{B*}). In summary, the positions of equilibria and their stability, and escalation or de-escalation are a function of the curves' position and shape. Figure 2 also illustrates how the convexity of the two response curves implies the existence of the unstable equilibrium in the northeast, motivating our interest in concavity and convexity of response curves (beyond intercepts and slopes).

Why would response curves have these shapes rather than others? Maximizing welfare subject to constraints might result in *convex* response curves which would better deter high levels of damage, by raising the marginal cost of inflicting increased damage for their opponent. One reason is a lack of capacity. Another is that disproportionate violence may be unacceptable to domestic or international audiences, making a convex response curve not credible. A third reason is the riskiness of accidentally landing to the right of the northeast equilibrium in Figure 2, doomed to escalation. Finally, the figure also illustrates that increasing convexity comes at the cost of making accidental escalation more likely. Imagine that A shifts to an even more convex response curve (though still anchored at the origin), by everywhere increasing the slope of the blue line. Doing so would shift the new northeast equilibrium to the west, narrowing the de-escalation interval, and *increasing the chance of an attack accidentally landing in the unbounded escalation interval* east of that equilibrium. Of course, that thought experiment assumes no strategic response by B, who might be scared into reducing convexity (or slope) –to prevent escalation, or react symmetrically –*further contracting the de-escalation space* (again assuming that B is not constrained by capacity, domestic or international pressure).

Key to our approach is the tension between short- and long-run damage suffered. Either side could immediately de-escalate to the low violence equilibrium, (at (0,0)), gaining relief from violence this round. Yet that would imply a flat response curve—which in the next round will fail to disincentivize attacks by the opponent (by not imposing costs). This inherent tension is known as the commitment problem in the literature on repeated games (e.g., Dixit and Skeath, pp. 292-306). The core implication is that sides seeking to maintain credibility will keep response curves stable between rounds. We will test for that stability empirically.

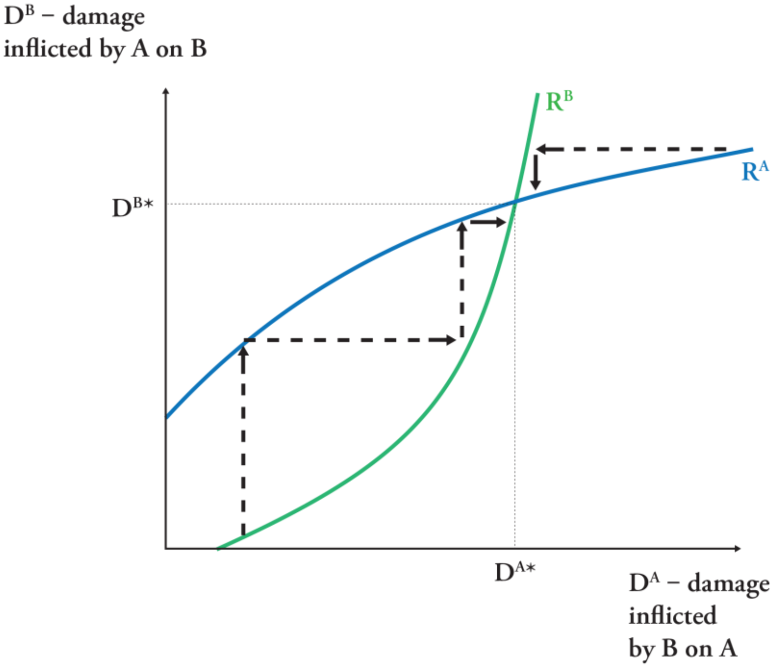
Relevant to escalation management are two further empirical questions: First, what are the upper bounds of the de-escalating interval (D^{A*}, D^{B*}), beyond which both sides are doomed to escalation?

Second, is the stable, low-violence equilibrium at (0,0)? In the literature on deterrence between two nuclear powers, the origin was an aspirational equilibrium of great importance, not least because de-escalating response curves at low levels of damage (presumably using conventional weapons) could not be assumed. So any deviation from the origin, which thankfully never occurred during the Cold War, might have been catastrophic.

In contrast, in the current discussion of integrated deterrence involving conventional weapons, we frequently see attacks and damage: for example, cyberattacks, election interference, sanctions, and other damaging violations of either international law or sovereignty. Attacks might occur in equilibrium depending on the shape of the response curves. A deterrent posture implies only that response curves are increasing in damage suffered without any other restriction on their shapes. This may lead to one stable equilibrium at the origin, but it could also lead to a stable equilibrium with positive violence (as we will estimate below), and even to multiple stable equilibria.

Figure 3 illustrates response curves that generate (a stable) equilibrium with incomplete deterrence, by which we mean that at least one side suffers (nonzero) damage in equilibrium. In this example, response curves are concave, and do not go through the origin. The vertical intercept of A's response curve is above the origin, and the horizontal intercept of B's response curve is to the right of the origin. This implies that A and B carry out "unprovoked" attacks, as they respond with positive damage even when they suffer zero damage. In this scenario, incomplete deterrence occurs: the resulting equilibrium involves damage to both sides. A secondary implication is that there is only one equilibrium.

Figure 3: Equilibrium with Incomplete Deterrence



We should emphasize that there is no completely peaceful equilibrium in this case. If the two sides find themselves at the origin, one or the other would initiate an unprovoked attack, leading to an escalating sequence that ends at the same violent equilibrium, as illustrated by the serrated lines southwest of the equilibrium. The result is stable (since the green line cuts the blue from below in the neighborhood of the equilibrium), though violent.

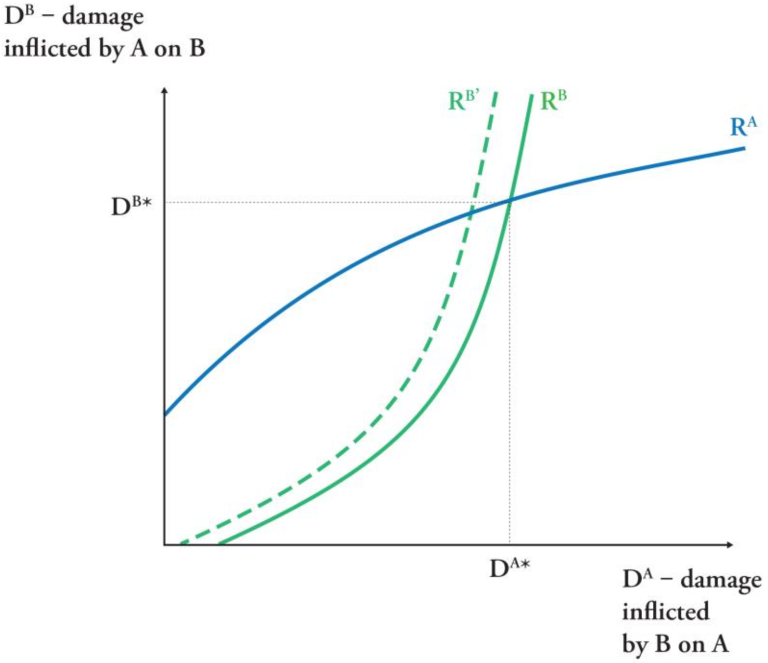
An incomplete-deterrence equilibrium invites the question of policy options. We turn next to actions the sides might take to shift such an equilibrium, by influencing an opponent's response curve.

2B. Deterrence by Denial, and by Resilience

The deterrence concept discussed so far operates by imposing costs –tracing an upward sloping response curve. It is distinct from deterrence by denial (the basic logic of conventional warfare) in which one attempts to reduce or degrade the opponent's capacity to attack (by destroying munitions, assassinating leaders with irreplaceable skills, or eroding public support for continued attacks, for example).

Figure 4 illustrates. Imagine that A destroys some capacity of B. The effect on B's response curve depends on B's underlying welfare and cost functions, which we are now forced to consider. A fairly general assumption is that destroying B's capacity will raise B's marginal cost of damaging A (more than it might increase B's marginal return from damaging A) at all levels of D^B. For instance, if Israel destroys some Gazan rockets with an airstrike, it may not fully preclude the possibility of a future high-damage attack, but could instead force Gazan militants to use more mortars (or put themselves at risk by moving mortars closer to the border fence) to achieve the same levels of damage. If so, A will have succeeded in shifting B's response curve leftwards to lower levels of damage, as shown by the serrated green line. The resulting equilibrium (to the bottom left, where the serrated response curve $R^{B'}$ meets R^A) will occur at lower levels of damage for both A and B, when compared to the previous equilibrium (though the welfare effects are presumably higher for A and surely lower for B).

Figure 4: Deterrence by Denial, and by Resilience



Our data actually include examples of three major Israeli military operations and other large actions that destroy Gazan capacity, so the hypothesis that they shift the Gazan response curve to a lower level of violence is testable. Those operations are often criticized as merely "mowing the lawn," in the sense that destroyed capacity might be restored (just as grass grows), yet in our model the operation might be nevertheless pursued by a rational actor if a long post-operation period of less violent equilibrium, even if limited, justifies the one-time cost of the operation (see also Inbar and Shamir 2014; Gibilisco 2023).

Deterrence by resilience, by contrast, reduces or even prevents damage associated with attacks initiated by an opponent. This should be the case for the "Iron Dome" air defense system, which arguably makes Israel more resilient, and may affect Gazans' damage function. Under the fairly weak assumption (about the underlying welfare function) that the increased resilience of A reduces the marginal benefit to B (through cost imposition) of attacks perpetrated at all levels of damage, we can predict how B's response curve will shift. In the diagram, if A increases resilience, that would shift B's response curve to lower levels of damage (to the left), as did deterrence by denial (above). The resulting equilibrium is again illustrated in Figure 4, at lower levels of net damage. As the diagram illustrates, increasing resilience has very similar effects on equilibrium outcomes as does destroying capacity –though it is achieved through defensive rather than offensive operations.

An additional prediction of a general strategic model is that a shift in an opponent's response curve may generate an adjustment of one's own. Under fairly general assumptions, both deterrence by denial and by resilience (shifts of an opponent's curve) will complement deterrence by cost-imposition (slope of one's own curve), so that a shift to the left of an opponent's response curve (i.e., to a less aggressive posture) will induce a shift to the left of one's own (more aggressive); the logic being that the benefits of cost-imposition remain the same, while the costs (in expected retaliation) have declined. A follow-on implication of that additional prediction is that major military operations (denial) and Iron Dome (resilience) unambiguously shift the subsequent equilibrium to lower damage for Israel. Whether the new equilibrium has lower damage for Gaza (as illustrated in the diagram) is theoretically ambiguous (because the blue response curve might shift as well).

Estimation will allow us to test the general applicability of this model in two important ways. To get an interval with de-escalation, the shape of response curves must yield a fixed point (i.e., a stable steady state) at the low end of the interval. Additionally, for these episodes to constitute a dynamic equilibrium as a group, the response curves must be *stable across episodes as well as within*. Note that the stability of response curves constitutes refuting evidence for a competing model, in which attacks reduce opponent capacity (deter by denial).

2C. Stochastic Response Curves

In our data responses are stochastic. The Israel Defense Forces (IDF) do not respond to 13 percent of initial attacks from Gaza during our sample period, while Gazan forces do not respond to 35 percent of initial Israeli attacks. Moreover, the intensity of response shows considerable variance, even conditional on damage suffered. That behavior motivates a model with mixed strategies—i.e., responding in probability, and with stochastic damage inflicted.

Why mixed strategies? Intuitively, mixed strategies would be preferable to always responding if there are fixed costs to responding at all, as a means of maintaining deterrence against relatively low levels of damage incurred (if marginal costs of response exceed marginal benefits). Mixed strategies also solve the dilemma of achieving deterrence by creating an expectation of costs imposed without being predictable about when (or how or where) a response will come, forcing the opponent to incur excess mitigation costs (e.g., taking shelter).

In the stochastic case, both A and B play mixed strategies. They either respond, with probabilities $P^A(D^A)$ and $P^B(D^B)$, respectively, or not. That is, both sides respond to damage (suffered) with an expected amount of damage inflicted. That approach is coherent if we generalize welfare maximization to expected welfare maximization assuming risk-averse opponents.

If the probability of response increases in damage suffered, and response curves (conditional on some response) are increasing in damage suffered, then expected response curves, such as $E(D^B) = P^A(D^A) \times R^A(D^A)$, will also increase in damage suffered. As in the deterministic case, the resulting expected response curves should arguably adopt a deterrence posture, but their exact shape is determined by the maximization of welfare subject to (political and capacity) constraints. The tension between short- and long-run returns to retaliation (described above) carries over to the stochastic case, as do the discussions of stability, de-escalation, and unique equilibria, though now in expected damage incurred, rather than deterministic damage.

The stochastic analysis is somewhat deceiving in its simplicity. An equilibrium in mixed strategies appears as a point in expected damage space on a graph, but actually represents a bivariate distribution of damage suffered by A and B concentrated around that point –which reflects only the *expected values* of damage.[7] Anticipating our empirical results, we will estimate a stable, stochastic equilibrium—the first we know of in this literature on conflict.

---

[7] Technically, those distributions are truncated at the horizontal and vertical axes, which is how we will implement them in estimation.

3. Data

Our study of mutual deterrence is enabled by the availability of a unique dataset describing in remarkable detail the repeated violent interactions between the IDF and Gazan militants. These data are unusual in their recording frequency, the richness of incident reports, and the availability of other supporting data, including from social media of the perpetrators, to validate and aid in interpretation and coding.

We build our main dataset using daily security reports of all attacks in Gaza recorded by the United Nations Field Security Office. The reports document attacks at up to five-minute precision, along with the governorate where they occurred.[8] We begin analyzing the UN reports on June 15th, 2007, the day Hamas seized control of the Gaza Strip from Fatah security forces. In our analysis sample we also omit the period after December 15 2017 when Gazans began regularly approaching the border fence, violating the buffer ("no go") zone that Israel claimed in order to secure the fence. While those events began as nonviolent, they eventually included acts of violence by young men, endorsed by all the militant groups, which drew Israeli warning shots, exchanges of live fire and often fatalities on the Gazan side.[9] We exclude them for two reasons: first, in retrospect those events constituted a contest over the buffer zone, which would be used to launch the attacks of October 7 2023; second, they represented a heightened willingness to take on risk by young men who are classified as civilians (rather than militants) in United Nations reports. Taken together, these events may reflect increased willingness by Hamas and other militant groups to incur casualties, or possibly a change in strategy, presenting a challenge that we defer to future research.

Figure 5 reproduces the security report of November 20th, 2010, a particularly violent day we've chosen to convey the richness of the data. Reports typically summarize an attack in two

---

[8] Gaza has five governorates. Attacks in two of them, North Gaza and Gaza, are reported together by the United Nations in the earlier years of our sample, so we combine those in our study.

[9] UN Daily Security Update of 15 December 2017 documents four demonstrations at separate locations adjacent to the security fence in three governorates, the same afternoon (of 14 December). Those demonstrations would prelude the "March of Return" confrontations on the border fence the following Spring. https://www.haaretz.com/israel-news/2018-03-31/ty-article/.premium/five-palestinians-reportedly-killed-by-israeli-army-as-thousands-rally-for-mass-gaza-protests/0000017f-e121-d7b2-a77f-e3277abf0000 downloaded 4/3/2024.

or three sentences. On November 19th, Gazan militants fired mortars at Southern Israel at 12:30 and 13:50; the Israeli Air Force (IAF) retaliated by firing 4 missiles at different targets in the Gaza Strip between 15:15 and 15:40; Gazan militants counter-retaliated by firing one mortar at 19:25, and the IAF attacked again at 23:30.

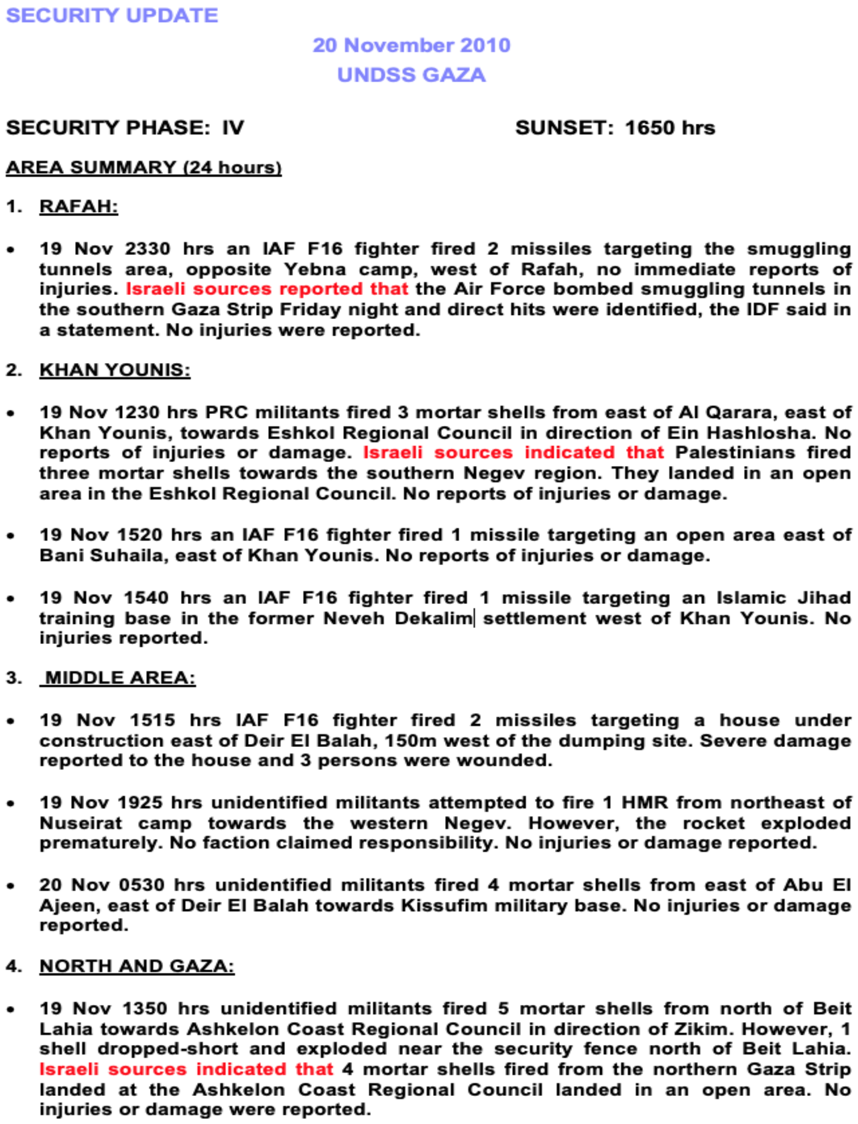


Figure 5: Security Report for November 20th, 2010

SECURITY UPDATE

20 November 2010
UNDSS GAZA

SECURITY PHASE: IV **SUNSET: 1650 hrs**

AREA SUMMARY (24 hours)

1. RAFAH:

- **19 Nov 2330 hrs an IAF F16 fighter fired 2 missiles targeting the smuggling tunnels area, opposite Yebna camp, west of Rafah, no immediate reports of injuries. Israeli sources reported that the Air Force bombed smuggling tunnels in the southern Gaza Strip Friday night and direct hits were identified, the IDF said in a statement. No injuries were reported.**

2. KHAN YOUNIS:

- **19 Nov 1230 hrs PRC militants fired 3 mortar shells from east of Al Qarara, east of Khan Younis, towards Eshkol Regional Council in direction of Ein Hashlosha. No reports of injuries or damage. Israeli sources indicated that Palestinians fired three mortar shells towards the southern Negev region. They landed in an open area in the Eshkol Regional Council. No reports of injuries or damage.**
- **19 Nov 1520 hrs an IAF F16 fighter fired 1 missile targeting an open area east of Bani Suhaila, east of Khan Younis. No reports of injuries or damage.**
- **19 Nov 1540 hrs an IAF F16 fighter fired 1 missile targeting an Islamic Jihad training base in the former Neveh Dekalim settlement west of Khan Younis. No injuries reported.**

3. MIDDLE AREA:

- **19 Nov 1515 hrs IAF F16 fighter fired 2 missiles targeting a house under construction east of Deir El Balah, 150m west of the dumping site. Severe damage reported to the house and 3 persons were wounded.**
- **19 Nov 1925 hrs unidentified militants attempted to fire 1 HMR from northeast of Nuseirat camp towards the western Negev. However, the rocket exploded prematurely. No faction claimed responsibility. No injuries or damage reported.**
- **20 Nov 0530 hrs unidentified militants fired 4 mortar shells from east of Abu El Ajeen, east of Deir El Balah towards Kissufim military base. No injuries or damage reported.**

4. NORTH AND GAZA:

- **19 Nov 1350 hrs unidentified militants fired 5 mortar shells from north of Beit Lahia towards Ashkelon Coast Regional Council in direction of Zikim. However, 1 shell dropped-short and exploded near the security fence north of Beit Lahia. Israeli sources indicated that 4 mortar shells fired from the northern Gaza Strip landed at the Ashkelon Coast Regional Council landed in an open area. No injuries or damage were reported.**

We record the attacker, target, type and quantity of munition, and casualties associated with each attack. For the Israeli side, this entails the number of airstrikes, small arms fire incidents, shelling, and incursions. For the Gaza side, reports include the number of small arms fire incidents, grad rockets, other rockets, and mortars. Deaths and injuries are recorded for Gazan civilians, Hamas operatives, other militants, and Israelis. The reports suggest that small arms fire is not always used to directly inflict damage. For example, the Israeli Defense Force would fire into the air at border fences to disperse large crowds, and the Israeli Navy would fire across the bow of Gazan fishing boats that drifted into restricted waters. We remove such situations by only coding Israeli small arms fire accompanied by an incursion or resulting in a casualty. (Our findings will be robust to including or excluding all Israeli small arms fire.)

Figure 6 presents daily tallies of Israeli and Gazan attacks between June 15th, 2007 and December 31st, 2018. In total, Israel launched 10,269 airstrikes, shelled Gaza 22,659 times, performed 1,379 incursions, and fired small arms 1,285 times (implementing our small arms inclusion rule discussed above). Gazan militants fired 10,207 rockets, 6,225 mortars, and 982 grad rockets into Israel. Gazan militants also fired small arms 355 times. The figure illustrates the high frequency of violence during our sample period. Out of 4,218 days included in the sample, Gazan militants performed some form of a violent attack in 33.1% of them (1,396 days) and Israel attacked in 45% of them (1,894 days). This violence caused 38 Israeli deaths (374 Israelis were injured). On the Gazan side we observe 3,522 fatalities, and 17,281 individuals were injured. Out of all Gazan fatalities, 748 of them are civilians and the remainder militants (2,070 injured Gazans are civilians).[11]

These numbers mask a great deal of variation in violence over time, as depicted in Figure 6. The figure shows a relatively low number of daily attacks, together with extremely violent periods occurring right before and during major confrontations between Israel and Hamas. These include the three main Israeli military operations during our sample period: Cast Lead (December 27th, 2008 until January 18th, 2009), Pillar of Defense (November 14th, 2012 until November 21st, 2012), and Protective Edge (June 12th, 2014 until August 26th, 2014). Gazan attacks are more evenly distributed over time, whereas Israeli attacks substantially increase during major operations. Note that violence substantially decreases after those operations vis-à-vis the level of violence before the operations. Within 30 days before those operations, Gazan militants attacked Israel 547 times and Israel attacked Gaza 267 times. The total number of

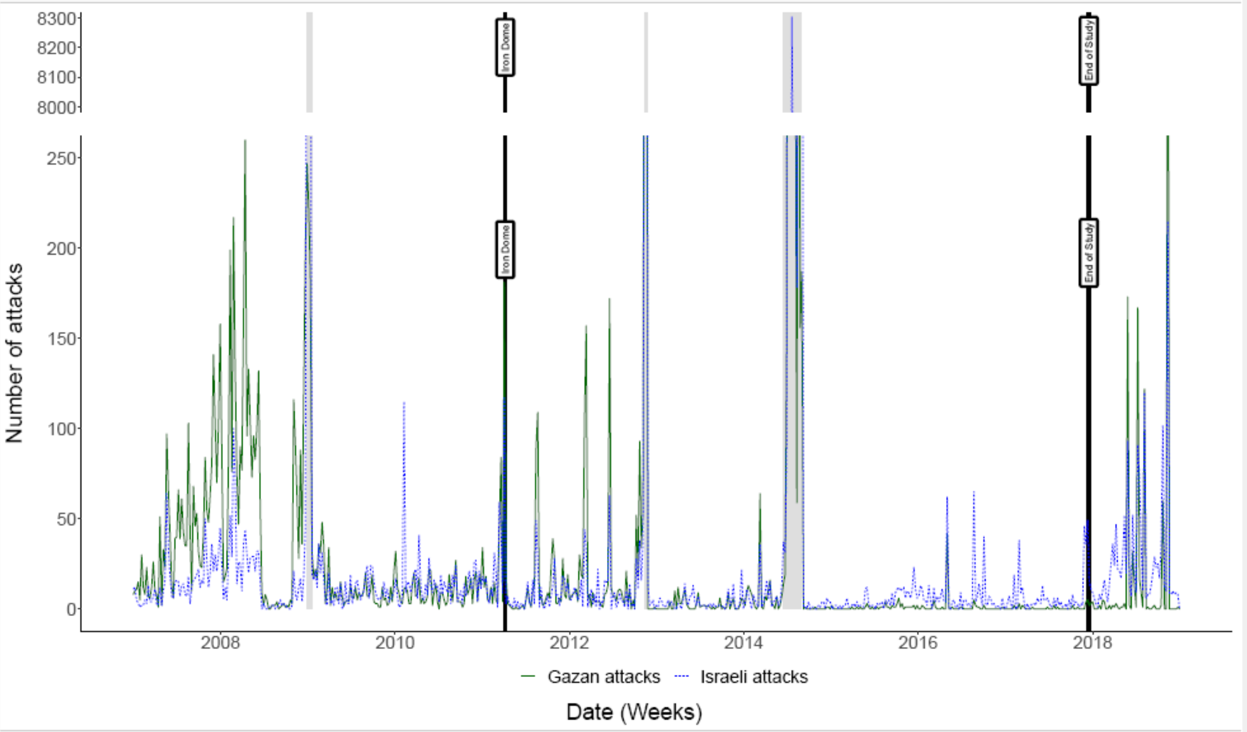
---

[11] Among militants, 227 fatalities were Hamas with the remainder belonging to other militant groups, including participants at violent protests at border fence (affiliation not listed, but protests may be organized by Hamas).

attacks decreases to 258 from Gaza and increases to 383 from Israel for the 30-day windows after those operations.

We omit the periods of the three major Israeli military operations in Gaza from our analysis sample for three reasons: First, descriptions of those operations characterize them not as escalatory responses to a particular attack (or sequence of attacks), but instead as pre-planned campaigns to reduce militant capacity. Indeed, the Israeli government announced that the purpose of operations in Gaza was to hamper Hamas' military capabilities, such as by demolishing tunnels or destroying munitions depots (Inbar and Shamir 2014). Second, Gazan militants' response (evident in Figure 6) would have been constrained by operational ability, being under attack. Third, even if those attacks and counterattacks fit on some response curve, their sequence of response and counter-response, and the available recording of that sequence, would have been disrupted by the very intensity of conflict around these operations.

Figure 6: Gazan and Israeli Attacks 2007-2018.



Note: Y Axis Break at 250 Attacks

4. Actions and Episodes

Even though the sides describe their attacks as retaliatory, they sometimes inflict a sequence of attacks before suffering a response. We address this by aggregating attacks into *actions*, where an action is a sequence of attacks executed by the same side with no attack executed by the other side between them.[13]

Actions, in turn, tend to occur in *episodes*, which we will think of as rounds in a repeated game. (We will provide evidence below that response curves are stable from episode to episode, reinforcing the idea that both sides treat these as rounds.) One challenge to coding is that the protagonists don't announce when an episode has ended, they simply stop retaliating. Our coding rule is to end an episode if there ensue 48 hours of *lull* (no attacks). We use a 48-hour lull rule to aggregate actions into episodes for two main reasons. First, the IDF and Gazan militants tend to retaliate within one or two days of the initial attack (as reported by their social media), with few exceptions.[14] Second, speedy retaliation makes sense because each side wants others to interpret its action as a retaliation, i.e., as the price of causing damage, establishing this causal linkage (in the minds of internal and external audiences). Moreover, both sides have the technological capability to retaliate within hours of an attack. Israel keeps fighter jets ready to scramble. It also has drones in the skies over Gaza and tanks around its borders which, according to our data, oftentimes spot and strike militant rocket crews even as they set up to launch. Similarly, as Haushofer et al. (2010) established, the technology for launching mortars and rockets allows Gazan militants to retaliate within hours of provocation.[15] An *episode* is thus defined as a group of actions bookended by lulls. (We also experimented with a *seven day* lull rule.) Recall that an episode ends when one side does not retaliate, which we code as a zero damage action. We code a lull as two zero damage actions for every 48 hours elapsed between (nonzero) actions.

---

[13] Marlin-Bennet et al (1991) refer to these as 'moves' and 'turns,' in an analogous coding.

[14] Abrahams et al (2019) report data on retaliation from the social media accounts of both the IDF and militants between April 2009 and July 2016, collected by the Meir Amit Intelligence and Terrorism Information Center (http://www.terrorism-info.org.il/). Of 305 retaliatory IDF airstrikes in that sample they find that at least 95% occurred within 2 days of initial attack. Of 26 retaliatory militant rocket attacks, 22 (85%) occurred within two days.

[15] This contrasts with suicide attacks launched from the West Bank during the Second Intifada, which potentially required weeks of planning (Jaeger and Paserman, 2008).

The UN records in which governorate attacks occurred, as we saw in Figure 5. We present three approaches to aggregate actions and episodes across governorates, each relying on different assumptions. The first is to assume that all actions in the Israeli-Gazan conflict are responses to the other side's actions, regardless of location (effectively assuming that Gaza is one conflict zone). We refer to this as the *broad* specification. A concern with this approach is measurement error in estimating response curves' slopes because we are likely estimating some actions as responses to damage suffered elsewhere, when they were not. For example, this aggregation may mistake an incursion in Rafah governorate as a response to mortar fire coming from Khan Yunis governorate, when these are independent attacks. Our second approach is to assume that actions are only responses to attacks within the same governorate. We refer to this as the *local* specification. This aggregation method may misrepresent actions responding across governorates as being unprovoked attacks beginning an episode. Econometrically, it might particularly bias the response curves' intercept because we are failing to attribute responses to actions outside the governorate. The final approach is a mix of local and broad specifications, which we refer to as *hybrid.* This combines episodes from the *local* specification together if an actor uses the same non-small arms munitions in two governorates within one hour of each other –our logic being that small arms fire may be controlled at the local level, whereas rules of engagement for more lethal munitions require approval at a higher level. We prefer this *hybrid* specification and use it throughout the remainder of the paper.[16] The results of the analysis (reported below) are robust to varying the 48-hour rule for lulls to one week, using the broad or local specification, linking hybrid episodes with small arms firings, and linking episodes using any munition fired by the same side within an hour of one another.

Table 1 reports on actions (omitting periods of major operations) for those three approaches to coding episodes. Focusing on the hybrid approach, Israel's attacks aggregate into 6,461 actions, and Gazans' attacks into 6,159. Importantly, Israel responds to only 33% (2,134/6,461) of Gazan actions, and Gazan militants respond to only 30%, which motivates our interest in mixed strategies. Most Israeli actions consist of shelling and airstrikes. They cause on average 0.13 fatalities and 0.52 injuries. Most Gazan actions consist of mortar and rocket

---

[16] See online appendix for a detailed explanation of the hybrid aggregation method.

attacks. They do not tend to cause Israeli fatalities but result in 0.04 injuries on average. The broad coding records less actions, because it aggregates more attacks across governorates into a single action, while the local coding disaggregates those same attacks into more actions. For that reason, broad episodes also record more fatalities and injuries, and local episodes less.

Table 1: Actions by Episode Coding - Means

	Local		**Hybrid**		**Broad**	
	Israel	**Gaza**	**Israel**	**Gaza**	**Israel**	**Gaza**
Number of Actions	8278	7843	6461	6159	2812	2612
Number of Episodes	1583	1583	1168	1168	489	489
Pr(Damage Caused>0)	0.29	0.25	0.33	0.30	0.68	0.66
Number of Nonzero Actions	2386	1951	2134	1832	1912	1712
E[Damage Caused|Damage>0]	0.42	0.04	0.47	0.04	0.53	0.05
	(0.54)	(0.05)	(0.69)	(0.06)	(0.73)	(0.06)
Airstrikes	0.24	-	0.31	-	0.71	-
	(1.13)	-	(1.37)	-	(1.98)	-
Shelling	0.28	-	0.36	-	0.83	-
	(2.28)	-	(2.84)	-	(4.27)	-
Incursion	0.14	-	0.18	-	0.41	-
	(0.41)	-	(0.5)	-	(0.77)	-
Small Arms Fire	0.1	0.04	0.13	0.05	0.29	0.11
	(0.37)	(0.2)	(0.5)	(0.22)	(0.76)	(0.33)
Mortars	-	0.51	-	0.65	-	1.53
	-	(2.38)	-	(2.79)	-	(4.16)
Rockets	-	0.51	-	0.65	-	1.54
	-	(1.94)	-	(2.15)	-	(3.01)
Grads	-	0.05	-	0.06	-	0.15
	-	(0.52)	-	(0.63)	-	(0.96)
Fatalities Caused	0.1	0.003	0.13	0.004	0.3	0.009
	(0.63)	(0.08)	(0.71)	(0.09)	(1.06)	(0.14)
Injuries Caused	0.4	0.03	0.52	0.04	1.18	0.1
	(2.24)	(0.82)	(2.87)	(0.93)	(4.3)	(1.42)

[a] Values are the average munition or casualty per each side's non-zero actions, with standard deviations in parentheses. Lull periods are included in calculating the averages. Episodes are determined using the 48 hour lull rule. Local episodes are coded as confined to a single governorate. Broad episodes are coded ignoring governorates. Hybrid episodes link local episodes if the same munition is fired within an hour in multiple governorates. Damage explained below.

Recall that we aggregate actions bookended by lulls into episodes, which we will treat as rounds of a repeated game. Table 2 reports characteristics of episodes.

Table 2: Characteristics of Episodes

	Local	**Hybrid**	**Broad**
Number of Episodes	1583	1168	489
A. Episode duration (number of days)			
Median	0.2	0.2	0.9
Mean	1.4	1.3	3.4
Max	120.8	121	130.1
Standard Deviation	4.2	4.6	8.7
B. Percent of episodes			
Begin with Gazan militant violence	41.1	42.1	34.6
With only Gazan militant violence	12	14	5
With only Israeli violence	31	32	29
C. Gazan militant actions per episode			
Minimum	0	0	0
Median	1	1	1
Mean	1.2	1.6	3.5
Max	106	248	260
Standard Deviation	3.5	8.3	14.5
D. Israeli actions per episode			
Minimum	0	0	0
Median	1	1	1
Mean	1.5	1.8	3.9
Max	105	247	260
Standard Deviation	3.4	8.2	14.4
E. Days without attacks between episodes (lulls)			
Minimum	2	2	2
Median	4.9	4.9	3.3
Mean	8.3	8.2	4.4
Max	103.8	92.4	80.3
Standard Deviation	10.5	9.5	4.4

[a] Calculations use the universe of firings, incursions, airstrikes, and projectile launches. Lull periods are included in calculating the averages. Episodes are determined using the 48 hour lull rule. Local episodes are coded as confined to a single governorate. Broad episodes are coded ignoring governorates. Hybrid episodes link local episodes if the same munition is fired within an hour in multiple governorates.

[b] UN Data from June 15, 2007 until December 15, 2017.

The main message emerging from the table is the variation in episode characteristics. Focusing on the hybrid episodic coding, duration ranges from five minutes (zero days) to 121 days, and from a single action up to almost 500 (248 Gazan + 247 Israeli). That said, most episodes are short-lived, with over half lasting less than five hours. The median lull length is 4.9 days. So a typical week might have 1-2 violent episodes during our decade long sample period.

Just less than half of episodes consist of attacks from only one side. Fourteen percent of episodes have only Gazan attacks and 32% have only Israeli attacks. Most of these single-action episodes consist of one attack (lasting for our minimum coded duration of five minutes), and do not cause any damage to the other side. Single-action episodes with Gazan attacks consist mostly of rocket attacks to open fields unretaliated by Israel. Single-action episodes composed of only Israeli attacks tend to be small arms fire during incursions.

Treating the conflict as a sequence of episodes and lulls will allow us to estimate response curves, and test their stability between episodes –a core assumption of the repeated game approach.

5. Measuring Damage

Estimating response curves requires a measure of predicted damage per attack that somehow combines different characteristics of harm, and associates them with different types of attacks. Characteristics range from the horrible to the more mundane, which in principle include not only fatalities but also physical injuries, damage to property, psychological trauma, as well as avoidance and mitigation costs. Fully measuring damage is impractical; we approximate using casualties (fatalities and injuries) reported after each attack, experimenting with different weights on fatalities and injuries.

We create an aggregate measure of predicted damage using as weights coefficients from linear regression of casualties on types of munition used by an opponent. Specifically, we estimate the following linear models to generate weights, where *i* denotes an attack.[17]

$$Damage\ to\ Gaza = \beta_1 small\ arms_i + \beta_2 incursions_i + \beta_3 small\ arms\ fire_i * incursions_i + \beta_4 shellings_i + \beta_5 airstrikes_i + \epsilon_i \tag{1}$$

---

[17] Recall that attacks are documented at five minutes intervals. A single attack often consists of a barrage using different types of munitions simultaneously.

$$Damage\ to\ Israel_i = \alpha_1 small\ arms\ fire_i + \alpha_2 mortars_i + \alpha_3 rockets_i + \alpha_4 grads_i + \epsilon_i \quad (2)$$

We allow no intercept, assuming that casualties cannot occur without an attack.[19]

Table 3: Predicted Damage by Type of Attack

	Predicted Damage to Gaza				**Predicted Damage to Israel**			
	Fatalities	**Fatalities + .2*Injuries**	**Fatalities + .5*Injuries**	**Fatalities + Injuries**	**Fatalities**	**Fatalities + .2*Injuries**	**Fatalities + .5*Injuries**	**Fatalities + Injuries**
Small Arms Fire	0.194***	0.665***	1.372***	2.550***	0.034***	0.050***	0.075***	0.116***
	(0.029)	(0.042)	(0.078)	(0.144)	(0.013)	(0.015)	(0.022)	(0.037)
Grads					0.002	0.034	0.082	0.163
					(0.001)	(0.022)	(0.056)	(0.112)
Rockets					0.006	0.013**	0.023***	0.040***
					(0.004)	(0.006)	(0.009)	(0.015)
Mortars					0.003	0.005	0.008*	0.013**
					(0.003)	(0.003)	(0.004)	(0.006)
Airstrikes	0.135***	0.217***	0.341***	0.546***				
	(0.017)	(0.026)	(0.040)	(0.065)				
Shellings	0.011***	0.019***	0.030**	0.049**				
	(0.004)	(0.007)	(0.012)	(0.020)				
Incursions	0.055***	0.084***	0.127***	0.199***				
	(0.016)	(0.024)	(0.037)	(0.059)				
Incursions x Small Arms Fire	−0.037	−0.463***	−1.102***	−2.167***				
	(0.056)	(0.075)	(0.119)	(0.202)				
Number of Attacks	3681	3681	3681	3681	3850	3850	3850	3850

[a] Expected damage is calculated using UN reports as the unit of analysis, referred to as attack-level data in the text. Every column i every panel presents OLS coefficients of a regression of a proxy for damage on weapons used. Heteroskedasticity robust standard errors appear in parentheses.

* p <.1, ** p<.05, *** p<.01

---

[19] All reported casualties are associated with at least one specific munition, so a statistically significant intercept would likely reflect a specification error in estimation.

Table 3 reports estimated coefficients for equations (1) and (2), exploring four alternative ways to measure damage using casualties. These are estimated using all attacks causing casualties (see the discussion of Table 1). Column 1 uses exclusively fatalities as our measure of damage; Column 2 assigns weight one to fatalities and 0.2 to injuries; Column 3 assigns weight one to fatalities and .5 to injuries; and Column 4 uses equal weights.

Estimated coefficients are quantitively proportional for different weighting schemes. All the coefficients are positive, as expected. Regardless of the weights used for casualties suffered by Gaza, small-arms fire has a higher coefficient than do incursions, airstrikes or shellings, with the relative size of coefficients similar regardless of the weight placed on injuries. Israeli small arms fire during incursions causes far less fatalities and injuries, so we estimate a separate coefficient for it, to achieve a more accurate weight. In terms of attacks suffered by Israel, small-arms fire again has the highest coefficient, followed by rockets and then mortars, regardless of the weight placed on injuries. The coefficient on Grad missiles is not statistically significant, perhaps because a large number of these missiles are intercepted by the Iron Dome system. (Though we do not find missiles to be any less lethal following the introduction of Iron Dome –in unreported results.) We return to discuss the Iron Dome missile defense below. Coefficients on small arms fire, rockets, airstrikes and incursions are quite precisely estimated.

Given that the results are qualitatively similar across weighting schemes, we use the fitted values from the second column (with a 0.2 weight on injuries relative to fatalities) as the preferred aggregate damage metric for each attack, as it seems most intuitively sensible. Recalling the discussion above of events contesting the border fence beginning December 2017, one reason to drop data from that period is the possibility that the Gazan militant weight on casualties had shifted, which would confound our measurement.

6. Estimating Response Curves

6A. Functional form and hypotheses

We choose simple functional forms to estimate response curves, to preserve precision in estimation.

$$P\,(D^{B}{}_{a} > 0) = P^{A} = \Phi(\alpha_0 + \alpha_1\, D^{A}{}_{a-1} + \alpha_2(D^{A}{}_{a-1})^{2}) \qquad (3)$$

is A's response probability, for action 'a', in response to damage suffered from 'a-1' by B, estimated as a quadratic Probit. Symmetrically,

$$P\,(D^{A}{}_{a} > 0) = P^{B} = \Phi(\beta_0 + \beta_1\, D^{B}{}_{a-1} + \beta_2(D^{B}{}_{a-1})^{2}) \qquad (4)$$

is B's response probability.

Conditional on nonzero response we estimate response intensity, which maps damage suffered into damage inflicted, as in Figure 1 of Section 2.

$$R^{A}{}_{a} \equiv D^{B}{}_{a} = \gamma_0 + \gamma_1\, D^{A}{}_{a-1} + \gamma_2\,(D^{A}{}_{a-1})^{2} + \varepsilon^{A}{}_{a} \mid D^{B} > 0\,, \qquad (5)$$

$$R^{B}{}_{a} \equiv D^{A}{}_{a} = \delta_0 + \delta_1\, D^{B}{}_{a-1} + \delta_2\,(D^{B}{}_{a-1})^{2} + \varepsilon^{B}{}_{a} \mid D^{A} > 0\,. \qquad (6)$$

Expected response curves, $E(D^{B}) = P^{A}(D^{A})\, R^{A}(D^{A})$, and $E(D^{A}) = P^{B}(D^{B})\, R^{B}(D^{B})$, (omitting subscript 'a' for readability) are quadratic in damage suffered to allow for concavity or convexity, recalling the importance of second derivatives (Section 2).

An attractive restriction on our expected response curves will be that they form a Tobit, that is the probability of response Probit and response intensity share the same coefficients, with a normal error (ε), so vectors of coefficients $\alpha = \gamma$ and $\beta = \delta$.

We test five related hypotheses:

1. Expected response curves $E(D^{B}) = P^{A}(D^{A})\, R^{A}(D^{A})$, and $E(D^{A}) = P^{B}(D^{B})\, R^{B}(D^{B})$, have positive slope, consistent with a deterrent posture (as in Figure 2).

2. Response curves are stable across episodes. We test by checking if lagged damage and lull length from past episodes (i.e., episodes e-j, for j>0) predict current damage.

3. Response curves are stable within episodes. We test by checking if lagged values of damage suffered preceding the last action by the opponent (a-1) predict current damage imposed (i.e., responses to action a-j for j>1).

4. The interaction of response curves yields a stable fixed point at nonnegative values of damage, or in other words, a stable equilibrium.
   Sufficient conditions are parameters of expected response curves that yield de-escalation to the right and escalation to the left. I.e., a fixed point,
   $R^B(R^A(D^A)) = D^A$ and $R^A(R^B(D^B)) = D^B$, with de-escalation
   $R^B(R^A(D^A)) < D^A$ and $R^A(R^B(D^B)) < D^B$, to the right, and escalation to the left (Figure 2).

   4A. If so, is that equilibrium at (0,0), which is to say, complete deterrence?
   The alternative is *incomplete deterrence*, with equilibrium at positive damage.

5. Major operations and the introduction of the Iron Dome defensive system do not shift the response curves of Gazan militants.
   Alternatively, major operations achieve deterrence, shifting Gazan militants' response curve to less violence (presumably by destroying capacity to impose damage on Israel – deterrence by denial), and the Iron Dome achieves deterrence (through resilience).

6B. Estimating Expected Response – Econometric Issues

Estimating response curves faces four econometric challenges particular to our approach. The data are a time-series, so we're sensitive to the possibility of serially correlated errors, which may reflect misspecification or omitted variables. Recall that our data are a sequence of actions, regardless of the duration of lags between actions. So any serial correlation in residuals must relate to previous residuals indexed by action (rather than calendar time). Second, actions come in episodes of various lengths, separated by lulls, and residuals could potentially be correlated between episodes as well as within. So we would like to test for serial correlation across episodes as well. Third, in the hybrid (and broad) coding of episodes, some episodes involve multiple governorates, and so we suspect higher variance in residuals, since those actions typically involve higher damage munitions and are more likely to involve multiple groups of militants. Finally, and most importantly, since actions alternate between sides and sides respond to damage suffered in the previous action, any correlation of ε^A_a with ε^B_{a-1}, implies inconsistent estimates of α and β, since $\varepsilon^{B}{}_{a-1}$ is correlated by construction with $D^{A}{}_{a-1}$ (should that residual have any variance at all).

Using the subscript 'e' to indicate an episode, we assume the following structure of possible error terms, recalling that a single lag indicates the action of the opponent, while two lags indicates the action of the same side.

$$\varepsilon^{A}{}_{e,a} = \rho^{AB}\,\varepsilon^{B}{}_{e,a-1} + \rho^{A}\,\varepsilon^{A}{}_{e,a-2} + \Theta^{A}\,\varepsilon^{A}{}_{e-1} + \eta^{A}\,1(.)_{e} + u^{A}{}_{e,a}\,, \qquad (7)$$
$$\varepsilon^{B}{}_{e,a} = \rho^{BA}\,\varepsilon^{A}{}_{e,a-1} + \rho^{B}\,\varepsilon^{B}{}_{e,a-2} + \Theta^{B}\,\varepsilon^{B}{}_{e-1} + \eta^{B}\,1(.)_{e} + u^{B}{}_{e,a}\,,$$

where $u^{A}{}_{e,a}$ and $u^{B}{}_{e,a}$ are independently distributed (and normally distributed to enable a Tobit). In this formulation, the cross-equation correlation is captured by ρ^{AB} and ρ^{BA}, the serial correlation over actions of the same side is captured by ρ^{A} and ρ^{B}, and the cross-episode correlations by Θ^{A} and Θ^{B}. The function 1(.) indicates a multi-governorate episode which multiplies independent normal random variables η^{A} and η^{B} with mean zero –capturing the possible extra variance in multi-governorate episodes.

Our challenges, then, are to find a way to calculate standard errors and properly sized tests in the presence of possible correlation between residuals in this unconventional time series of residuals. This is particularly important to us, given the centrality of testing Hypotheses 1-5 to our objective of modelling deterrence.

Our main approach will be to model this heteroskedasticity explicitly in estimating response curves [(3),(4),(5) and (6)]. We will also allow cluster-robust standard errors within episodes for testing.[20]

---

[20] We have also tried testing for serial correlation by expanding on a method suggested by Wooldridge (2010) for Tobit estimators. Additionally, we've estimated the parameters of (7) using the resulting Tobit residuals, allowing cluster-robust residual variance. We then simulate using estimates from (7) and our Tobit estimates, and compare results to estimate bias due to nonzero cross-correlations. We may eventually report all this in an appendix along with results of an alternative approach following MacKinnon et al (2023).

6C. Estimating Expected Response – Data

In estimating response curves the unit of observation is an action (which may include multiple attacks, as described above). Table 4 reports descriptive statistics for 6915 Israeli actions and 6531 actions by Gazan militants recorded during our sample period.

Table 4: Summary Statistics: Actions

Outcome	Number of Actions	Mean	Std. Dev.	Median	Min	Max
Israeli Response						
Damage to Gaza	6461	0.160	0.45	0.00	0.00	8.23
Pr(Damage to Gaza>0)	6461	0.330	0.47	0.00	0.00	1.00
Damage to Gaza|Damage>0	2136	0.470	0.69	0.22	0.02	8.23
Lag(Damage to Israel)	6461	0.010	0.04	0.00	0.00	0.54
Lag(Damage to Israel^2)	6461	0.001	0.01	0.00	0.00	0.30
Lull Length	6461	231.6	292.38	120.25	48.00	2218.33
Damage to Gaza Past Episode	6461	0.100	0.49	0.01	0.00	16.22
Damage to Gaza Past Episode^2	6461	0.250	4.75	0.00	0.00	263.11
Damage to Israel Past Episode	6461	0.940	3.35	0.08	0.00	83.80
Damage to Israel Past Episode^2	6461	12.110	144.32	0.01	0.00	7022.62
Gazan Response						
Damage to Israel	6159	0.010	0.04	0.00	0.00	0.54
Pr(Damage to Israel>0)	6159	0.300	0.46	0.00	0.00	1.00
Damage to Israel|Damage>0	1833	0.040	0.06	0.02	0.00	0.54
Lag(Damage to Gaza)	6159	0.160	0.46	0.00	0.00	8.23
Lag(Damage to Gaza^2)	6159	0.240	1.90	0.00	0.00	67.81
Lull Length	6159	230.1	293.58	118.5	48	2218.33
Damage to Gaza Past Episode	6159	0.100	0.54	0.01	0.00	16.22
Damage to Gaza Past Episode^2	6159	0.300	5.90	0.00	0.00	263.11
Damage to Israel Past Episode	6159	0.950	3.52	0.08	0.00	83.80
Damage to Israel Past Episode^2	6159	13.280	171.08	0.01	0.00	7022.62

[a] Summary statistics exclude the three major operations. See Section 3 for details

[b] The table provides summary statistics for actions. An action is a group of attacks uninterrupted by the other side and with less than 48 hours between attacks. A group of actions with less than 48 hours between their start times is an episode. See section 4 for more details. Damage is predicted, using fatalities + 0.2* injuries, following discussion of Table 3 above. Lull length measured in hours.

Beginning with Israeli responses, within episodes Israel often takes no action at all; the probability of response is 0.33. When it does, average damage inflicted by Israel per attack is 0.47, where damage is measured as predicted damage associated with the munitions used in an attack (using as weights the coefficients estimated in Table 3, column 2, so that damage is the equivalent of one fatality or 5 injuries). We choose predicted damage rather than actual damage suffered, following the logic of a response to *intent* rather than a response to *outcome*, and to capture disruption and mitigation costs in addition to actual casualties (for example, anxiety and moving to shelter).

As right-hand side variables we include damage to Israel in the previous action (including zero damage actions) and its square. Note that median damage suffered is zero, because a typical action is unprovoked –as will be true for Gazan response as well. To check the basic modeling assumption of the repeated game, that episodes are self-contained, we also include the length of the lull preceding the previous episode (measured in hours), damage to Israel in the previous episode, and damage to Gaza (inflicted by Israel) in the previous episode. Those could be motivated by the logic of a delayed response. Note that minimum damage in the previous episode is zero, which occurs when an unprovoked action is not responded to.

Symmetrically, Gazan militants record 6159 actions with average damage inflicted of 0.04 (fatality-equivalents). Gazan militants respond to Israeli actions 31% of the time. Damage to Gaza is much greater than that inflicted on Israel, reflecting the asymmetric capabilities of the two sides.

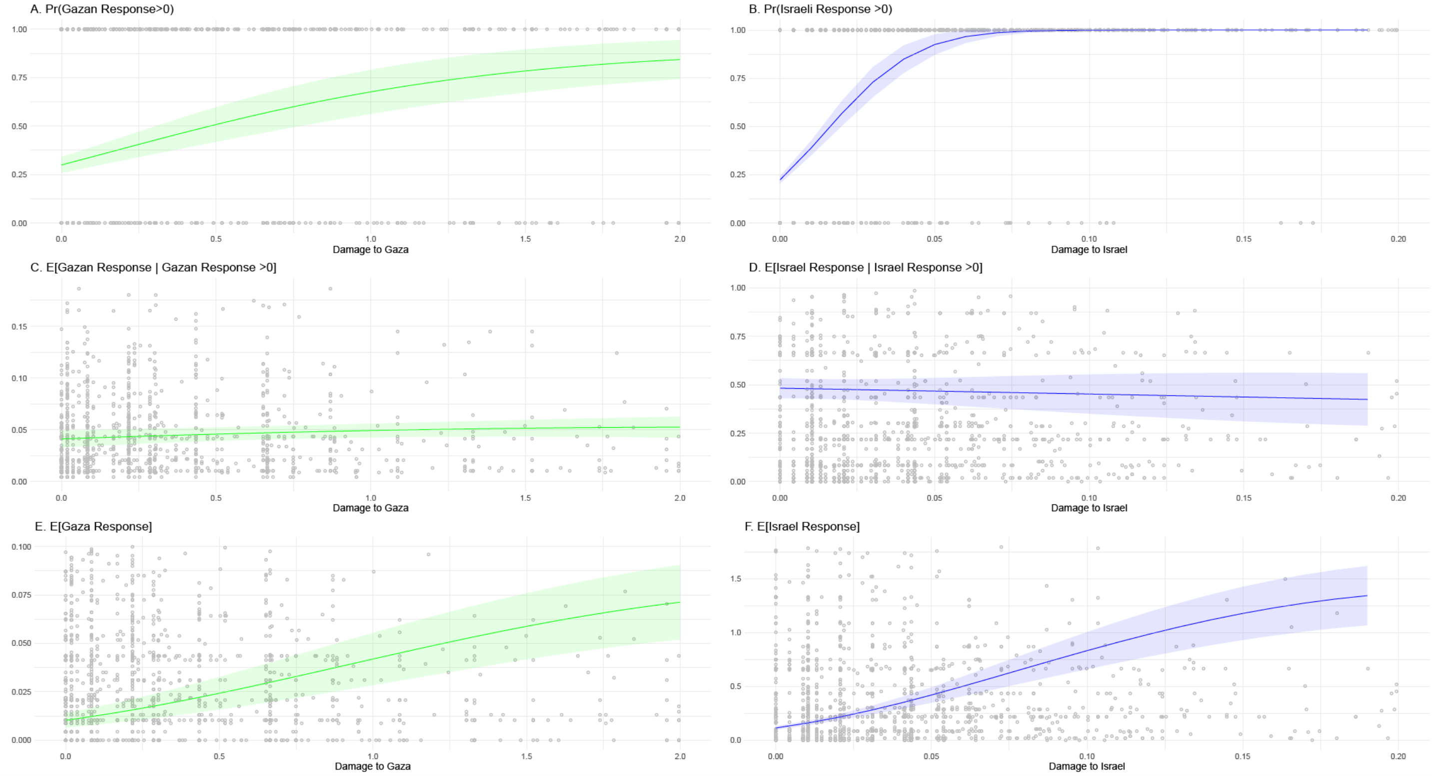
6D. Expected Response Curve Estimates

We estimate both probability of response and damage, as functions of damage suffered, in order to calculate expected response curves. Figure 7 illustrates our approach, with dots indicating data. Panels A and B display the Gazan and Israeli probabilities of response, respectively, estimated using quadratic probit curves, (again, using casualties to measure predicted damage on the x-axis calculated using coefficients from Table 3, column 2). Both Israeli and Gazan probabilities of response are increasing, though more so the Israeli.

Panels C and D graph damage caused in a single action on the y axis, against damage suffered on the X axis, for actions with nonzero damage caused. Plotted lines are from quadratic Tobit estimates, reported in Tables 5 and 6 below (with average marginal effects presented here). The Gazan line is mildly upward sloping with the Israeli line slightly downward sloping.

Panels E and F illustrate the implied expected response functions for Gaza and Israel, respectively, with the lines plotted for Tobit estimates reported below (in Tables 5 and 6). Both are upward-sloping in damage suffered, indicating a deterrent posture.

Figure 7: Estimating Expected Response



Tables 5 and 6 describe response curve estimates. Beginning with the Israeli in Table 5, the leftmost column reports OLS regression of Israeli damage inflicted on damage suffered. That specification cannot be literally correct, since the probability of response is inherently nonlinear, but it provides a simple description of what we will see in the other specifications. The positive coefficient on the linear term (5.42) combined with the negative coefficient on the quadratic term (-12.4) indicate an upward-sloping, concave expected response curve. The positive intercept of 0.11 indicates a propensity to inflict damage when none is suffered. Those patterns will persist in more complex specifications: a deterrent posture and an inclination to attack when no provocation is recorded, as we see in Figure 7.

Table 5: Israeli Response Curve

	OLS	Probit	Tobit						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Lag(Damage to Israel)	**5.42*****	**51.9*****	**18.3*****	**15.1*****	**12.4*****	**12.1*****	**11.1*****	**10.4*****	**10.3*****
	(0.73)	**(5.5)**	**(2.4)**	**(1.4)**	**(1.8)**	**(1.8)**	**(1.8)**	**(1.7)**	**(1.7)**
Lag(Damage to Israel^2)	**−12.4*****	**−94.6*****	**−40.6*****	**−33.6*****	**−23.6*****	**−23.1*****	**−22.1*****	**−20.4*****	**−20.3*****
	(2.1)	**(8.9)**	**(7.0)**	**(4.6)**	**(4.6)**	**(4.9)**	**(4.8)**	**(4.3)**	**(4.3)**
Lag(Damage to Israel,2)							3.20***	2.7***	3.04***
							(0.97)	(1.0)	(0.77)
Lag(Damage to Israel^2, 2)							−5.8**	−4.6*	−5.2**
							(2.5)	(2.5)	(2.2)
Lull length								−0.00035***	−3.6e−04***
								(0.00006)	(6.2e−05)
Damage to Israel Past Episode									−0.073
									(0.093)
Damage to Gaza Past Episode									0.0087
									(0.0179)
Intercept	0.1065***	−0.873***	−0.706***	−0.527***	−0.52***	−0.50***	−0.524***	−0.439***	−0.438***
	(0.0074)	(0.032)	(0.098)	(0.049)	(0.05)	(0.05)	(0.046)	(0.047)	(0.049)
log(σ)			−0.096	−0.543***	−0.539***	−0.58***	−0.564***	−0.561***	−0.554***
			(0.085)	(0.056)	(0.055)	(0.06)	(0.056)	(0.054)	(0.054)
I(Multi-Govenorate Episode)				0.617***	0.696***	0.692***	0.656***	0.656***	0.665***
				(0.075)	(0.077)	(0.076)	(0.082)	(0.086)	(0.087)
Lag(Damage to Israel)					−2.16***	−2.0**	−2.0**	−2.19***	−2.2***
					(0.73)	(0.8)	(0.8)	(0.81)	(0.8)
Damage to Gaza Past						0.055*	0.061*	0.059*	0.050
						(0.032)	(0.034)	(0.034)	(0.031)
Damage to Israel Past						−0.36**	−0.37**	−0.37**	−0.38**
						(0.15)	(0.16)	(0.16)	(0.16)
Number of Actions	6461	6461	6461	6461	6461	6461	6461	6461	6461
AIC	7754.0	6082.2	9164.8	8639.1	8595.9	8510.7	8480.5	8414.3	8404.0
BIC	7774.4	6102.5	9164.8	8639.1	8595.9	8510.7	8480.5	8414.3	8404.0
Number of Episodes	1170	1170	1170	1170	1170	1170	1170	1170	1170
Mean Outcome	0.16	0.33	0.16	0.16	0.16	0.16	0.16	0.16	0.16
R2/Psuedo R2	0.06	0.35	0.04	0.08	0.09	0.1	0.09	0.09	0.1

[a] Location estimates refer to the estimate of the mean while scale estimates are the estimates for the parameterized variance.

[b] Column 1 shows a simple OLS estimation of the reaction curve. Column 2 shows the coefficients for the latent variable from a Probit specification. Columns 3-11 show the coefficients for the latent variable from a Tobit specification. [c] Actions are aggregated across governorates if the same munition is fired by one side within an hour of one another. [d] There are six control variables: 1) the lull length prior to two previous episodes, 2) the total damage Israel incurred last episode, 3) the square of the total damage Israel incurred last episode, 4) the total damage Gaza incurred last episode, 5) the square of the total damage Gaza incurred last episode, and 6) an indicator if the episode was aggregated across governorates. [e] Following Wooldridge, the R2 for the Probit and Tobit specifications is the square of the correlation coefficient between the observed outcome and predicted outcome. [f] Major operations indicator includes an indicator for time periods between each major operation, as well as an indicator before and after Iron Dome was introduced. [g] Standard errors are clustered at the episode level using the 48 hour lull rule.

* p < 0.1, ** p < 0.05, *** p < 0.01

The Probit curve (in the next column to the right) has a similar shape, but is confined to predicting inside the [0,1] interval (as in Figure 7, on the top right, where the positive intercept is evident). The Tobit (column (3)) estimates the propensity (Probit) and intensity equations with common coefficients,[22] indicating the upward-sloping expected response curve in the bottom right panel of Figure 7.

Our residuals may be complex: correlated between sides, across actions, and over periods, and perhaps having extra variance when responses cross governorates (in the hybrid coding of episodes), as reflected in equation (7) above. To account for any possible resulting heteroskedasticity we estimate a more flexible specification in which the log standard error (σ) is a function of covariates, including an indicator for multiple governorates (column 4), and adding in lagged damage suffered by Israel (column 5), and damage to both side in the previous episode (column 6). All of those variables predict statistically significant changes to sigma, providing strong evidence that the Tobit specification displays heteroskedasticity, and indirect evidence that the coefficients in equation (7) are nonzero. Those heteroskedasticity accommodations shrink the linear and quadratic coefficients as well as the intercept of the response curve (comparing columns (3) and (6)) but it seems to remain upward sloping and concave, with coefficients more precisely estimated.

We turn now to whether the data fit a model of repeated games, beginning with whether response curves are stable within episodes (section 6A). I.e., is the response to damage just suffered in the last action, or to previous damage from actions within the same episode? Column (7) adds lagged damage within episode from the opponent's action before the last (zero damage if that action before the last was during a lull) using both linear and quadratic terms. This test yields a mixed result: coefficients on damage to Israel from opponent's action before the last (action a-3) look like shrunken versions of coefficients to opponent's last action (a-1), they have the same signs, are about a third to a quarter the size, and are statistically

---

[22] Note that Probit and Tobit coefficients are not directly comparable as reported, since the former has variance calibrated to one (as it is not identified in the standard specification). Comparison allows a simple specification check (Wooldridge, 2010), which we can do by dividing Tobit coefficients (column 3) by estimated standard deviation of the Tobit e^(-.096) = .91. Our estimates in columns (2) and (3) clearly fail that test, which further motivates heteroskedasticity adjustments and a longer regression in the remaining columns.

significant. So the simple model is rejected, a point we return to when analyzing equilibrium behavior below.

Another key assumption is that previous rounds are irrelevant, which is to say response curves are stable across episodes. Columns 8 and 9 add variables reflecting information from previous rounds, lull length preceding the previous episode,[23] damage incurred in the previous episode, and damage inflicted in the previous episode. The result here is mixed. On the one hand, data from the previous episode is predictive, and in intuitive ways. A long lull predicts a smaller (less aggressive) intercept, while damage suffered and damage inflicted in the previous episode are both statistically significant predictors (though hard to interpret as they must be collinear). On the other, estimated coefficients are small relative to those in the first two rows, and adding them hardly changes the slopes or intercept (comparing columns (7) and (9)). So a simple model based only on information from the current episode can be rejected statistically, but seems to provide a fairly good approximation of the Israeli response, and shows little sign of omitted variable bias when the specification is expanded.

Table 6 allows us to test the same three hypotheses for the response curve of Gazan militants. We follow the same sequence of specifications as in Table 5. The results differ mostly in finding even less within-episode stability. The leftmost column reports on an OLS regression of Gazan damage inflicted on damage suffered by Gazans, encompasses both extensive and intensive margins. That specification cannot literally be accurate, but captures the curvature of the response function which we will see in the Tobit: it reflects deterrence, sloping upwards, and is concave.[24] The positive intercept of 0.0096 indicates a low but positive propensity to inflict damage when none is suffered. The Probit curve (in the next column to the right) is more clearly concave (Figure 7, on the left). The Gazan Tobit estimates, in the next column to the right, are those that produce the concave curve in the bottom left of Figure 7.[25]

---

[23] The lull length preceding the *current* episodes is endogenous by construction, since it ends with an action.

[24] This may reflect a faulty measure of how Gazan militants, especially Hamas, perceive damage suffered to civilians. In an extension we plan to use more detailed data on the identity of Gazan casualties rather than assuming symmetry in how the two sides treat casualties. See Section 7 for details.

[25] Probit and Tobit coefficients are not directly comparable as reported, since the former has variance calibrated to one. Comparison allows a simple specification check, as Wooldridge (2010) , which requires dividing Tobit coefficients (column 3) by estimated standard deviation of the Tobit e^(-2.518) = .081. Our estimates in columns (2) and (3) actually pass that test (1.5 vs. .098/.081 = 1.21; -.027 vs. -.018/.081 = -0.23).

To account for possible heteroskedasticity due to the complex residual reflected in equation (7) above we allow a flexible specification in which the log standard error (σ) is a function of covariates, including an indicator for multiple governorates (column 4), and adding in lagged damage suffered by Gaza (column 5), and damage to both sides in the previous episode (column 6). As in Table 5, episodes including responses to actions in other governorates have higher variance, and damage from past episodes predicts variance. So the Gazan response curve shows heteroskedasticity, which again provides indirect evidence that coefficients in equation (7) are nonzero. These heteroskedasticity adjustments shrink estimated linear and quadratic coefficients in the first two rows, as well as the intercept (comparing columns (3) and (6)) but the Gazan estimated response curve remains upward sloping and concave, with slightly less precise estimates.

The results in the next column test whether Gazan militant response curves are stable *within* episode by adding to the right-hand side damage suffered in the action before the previous (a-3), in the same episode (of at the tail end of the last lull). This is a humbling moment for our simple model, as we find coefficients of about the same magnitude as those in response to the last action (a-1), also showing concavity and also statistically significant. Which is to say damage suffered by Gaza by action a-3 generated a response in action a-2, and approximately the same expected response *again* in action a. We discuss implications below.

Finally, we add information from previous episodes, to test the second hypothesis of Section 6A, that response curves are stable across episodes. Lull length before the previous episode is statistically significant, resulting in a less negative intercept but very little change in slope coefficients (in the first and second rows). Adding damage from the previous episode produces statistically significant response coefficients of similar size to those to damage just sustained (comparing columns (7) and (9) in the first through fourth rows). In sum, there is evidence for response curves shifting between episodes, but a specification ignoring past episodes remains a very good approximation.

Table 6: Gazan Response Curve

	OLS	Probit	Tobit							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Lag(Damage to Gaza)	**0.0280*****	**1.5*****	**0.0981*****	**0.0427*****	**0.0404*****	**0.0391*****	**0.0446*****	**0.038*****	**0.0364*****	**0.139*****
	(0.0037)	**(0.3)**	**(0.0047)**	**(0.0038)**	**(0.0047)**	**(0.0048)**	**(0.0057)**	**(0.005)**	**(0.0032)**	**(0.012)**
Lag(Damage to Gaza^2)	**−0.005*****	**−0.27****	**−0.0184*****	**−0.0097*****	**−0.0084*****	**−0.0084*****	**−0.0120*****	**−0.0096*****	**−0.00567*****	**−0.0446*****
	(0.001)	**(0.14)**	**(0.0013)**	**(0.0012)**	**(0.0019)**	**(0.0020)**	**(0.0028)**	**(0.0022)**	**(0.00081)**	**(0.0079)**
Lag(Damage to Gaza,2)							0.0301***	0.026***	0.0317***	0.0217***
							(0.0041)	(0.004)	(0.0038)	(0.0038)
Lag(Damage to Gaza^2,2)							−0.0053***	−0.0043***	−0.00395***	−0.0040***
							(0.0013)	(0.0012)	(0.00092)	(0.0011)
Lag(Damage to Gaza,4)										0.0312***
										(0.0037)
Lag(Damage to Gaza^2,4)										−0.0054***
										(0.0011)
Lull Length								−4.6e−05***	−4.6e−05***	−3.9e−05***
								(4.5e−06)	(4.6e−06)	(4.2e−06)
Damage to Israel Past									0.0449***	0.0174***
									(0.0077)	(0.0038)
Damage to Gaza Episode									−0.0057***	−0.00167***
									(0.0015)	(0.00037)
Intercept	0.0096***	−0.742***	−0.0617***	−0.0370***	−0.0370***	−0.0367***	−0.0403***	−0.0299***	−0.0307***	−0.0251***
	(0.0019)	(0.078)	(0.0021)	(0.0014)	(0.0014)	(0.0014)	(0.0016)	(0.0016)	(0.0016)	(0.0015)
Log (σ)			−2.518***	−3.289***	−3.286***	−3.284***	−3.257***	−3.260***	−3.225***	−3.35***
			(0.018)	(0.027)	(0.027)	(0.027)	(0.028)	(0.028)	(0.029)	(0.03)
I(Multi-Govenorate				0.958***	0.967***	0.963***	0.879***	0.878***	0.866***	0.744***
				(0.031)	(0.033)	(0.033)	(0.036)	(0.036)	(0.035)	(0.036)
Lag(Damage to Israel)					−0.030	−0.023	0.038	0.0074	−0.159***	0.242***
					(0.041)	(0.041)	(0.044)	(0.0408)	(0.045)	(0.044)
Damage to Gaza Past						−0.0296***	−0.0305***	−0.0321***	−0.015	−0.067***
						(0.0085)	(0.0091)	(0.0091)	(0.018)	(0.013)
Damage to Israel Past						0.195***	0.199***	0.200***	0.155	0.373***
						(0.053)	(0.056)	(0.056)	(0.098)	(0.077)
Number of Actions	6159	6159	6159	6159	6159	6159	6159	6159	6159	4991
AIC	−23664.0	6910.2	−322.1	−1271.6	−1270.1	−1280.2	−1354.1	−1501.9	−1563.1	−2770.0
BIC	−23643.8	6930.3	−295.2	−1238.0	−1229.8	−1226.4	−1286.8	−1427.9	−1475.7	−2672.3
Number of Episodes	1170	1170	1170	1170	1170	1170	1170	1170	1170	1005
Mean Outcome	0.01	0.3	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
R2/Psuedo R2	0.043	0.107	0.03	0.165	0.165	0.17	0.176	0.181	0.154	0.179

[a] See notes to Table 5. * p < 0.1, ** p < 0.05, *** p < 0.01

Returning to the major rejection of the simple repeated game model, the instability of response curves within episodes, we experiment with adding an additional lag of damage to Gaza (linear and quadratic) in column 10. The two coefficients are again statistically significant, though much smaller, at about a third the size of those in the first two rows. This test comes with a grain of salt, because the sample size shrinks by 19% when we ask for coefficients on actions with four lags.

In the following section we describe first equilibrium and dynamics in the simple model, then discuss the implications of responses to previous lags within episode for equilibrium and dynamics in the more complex model that our estimates suggest.

6E. Equilibrium

We begin by describing equilibrium in the simple model, in which reaction is only to the last action, then proceed to the more complex model (with extra lagged action coefficients) that the data demand in Tables 5 and 6. The top panel of Figure 8 describes the equilibrium predicted by simple estimated expected response curves of the IDF and Gazan militants, taken together, using the Tobit specifications from Tables 5 and 6 (column 8, coefficients in first two rows). The Gazan expected response curve, in green, has a positive intercept and a slightly concave positive slope. The Israeli expected response in blue, also has a positive intercept and a slightly concave positive slope

Since both curves are concave, they intersect only once. Because the Gazan line cuts the Israeli from below, that single intersection is a stable equilibrium, de-escalating from the right (i.e., from higher levels of violence), and escalating from the left, as in Figure 3. So the mutual deterrence equilibrium in this simple model results in a single, stable equilibrium, ruling out the possibility of the type of escalation illustrated in Figure 2.

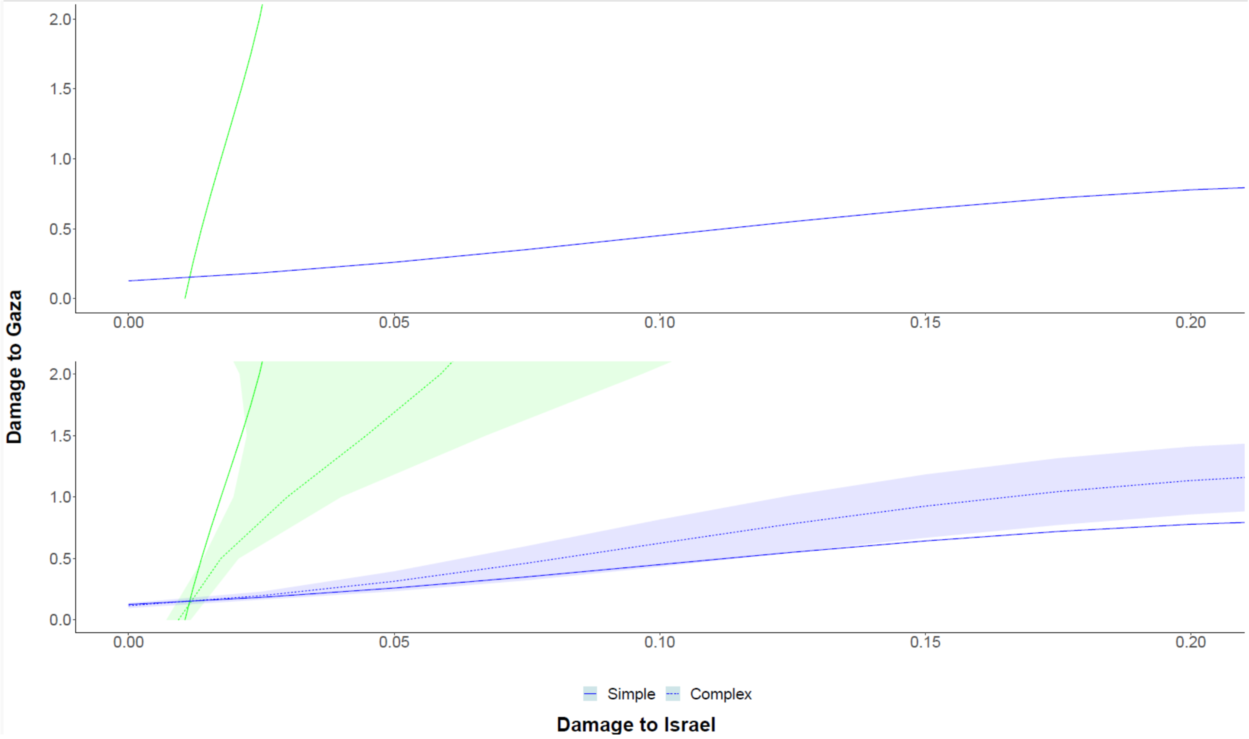


Figure 8: Equilibrium with Incomplete Deterrence: Simple and Complex Models

The less positive conclusion of this analysis is that the origin is an unstable point. Both intercepts are nonzero, so that zero violence cannot be sustained to the southwest of the equilibrium. E.g., zero violence incurred in Gaza induces a response at the intercept of the green line, which in turn induces an Israeli response, at least in probability. In this sense, this stable equilibrium characterizes *incomplete deterrence,* as illustrated in Figure 3. The estimated equilibrium location in the top panel is (0.015, 0.15), in damage units, –which measure expected fatalities and injuries.

To our knowledge, this is the first empirical characterization of a stable deterrence relationship using estimated response curves within episodes of conflict. Yet the diagram is deceivingly simple: equilibrium is not a point but a joint distribution reflecting the stochastic response of both sides (centered around the point of intersection). Attacks and actions will generate sequences of damage (within our episodes) that return to this equilibrium in probability, but not deterministically.

Yet that simple model is rejected by the data, in the sense that damage suffered in lagged actions previous to the last also predicts response, as we see in rows three and four of Tables 5 and 6 (columns 7-10). That more complex model creates a four-dimensional problem, which wouldn't lend itself to the graphical analysis of Figure 8.

Fortunately, analysis of an equilibrium simplifies. At a *stable* fixed point (D^{Af}, D^{Bf}), expected reaction curves $R^B(R^A(D^{Af})) = D^{Af}$, $R^A(R^B(D^{Bf})) = D^{Bf}$ have $(D^{Af}_{a-1}, D^{Bf}_{a-2}) = (D^{Af}_{a-3}, D^{Bf}_{a-4})$. So we can graph candidate stable damage combinations against each other using all the estimated lag action coefficients (Column (9), first four rows, Tables 5 and 6), to solve equilibrium and test for stability. The result is the two dotted curves in the bottom panel of Figure 8.

Those two dotted curves have a single intersection point, which is stable (the green line cuts the blue from below). That's a relief, as it makes the exercise internally coherent. They display a more deterrent posture than does the solid line, as their slope includes two actions of damage imposed for each of damage suffered. The equilibrium their crossing point describes is again one of incomplete deterrence, in the sense that there is violence at equilibrium. It happens to also be close enough to the origin, that equilibrium in the complex model is indistinguishable from that of the simple model. In sum, though the data insist on a complex model, and it indeed would have much trickier dynamics, the implication for equilibrium location is negligible.

What about the implication of between-episode instability? Again, though the data reject the null hypothesis that information from previous episodes is irrelevant, the size of the estimated coefficients indicates very little influence on the location of the equilibrium illustrated in Figure 8.

In summary, the workhorse simple model of repeated games oversimplifies the nature of this conflict, but provides an excellent approximation of the behavior of both sides. Though damage suffered in the last action of the current round can be rejected as a sufficient statistic, analysis based on the last action only can predict quite well the magnitude of responses, equilibrium uniqueness, equilibrium stability, escalation or de-escalation, and equilibrium location. For practical purposes, the model works extremely well.

As an application of that model, we turn to analysis of four major events, three Israeli operations in Gaza and the introduction of the Iron Dome missile defense system.

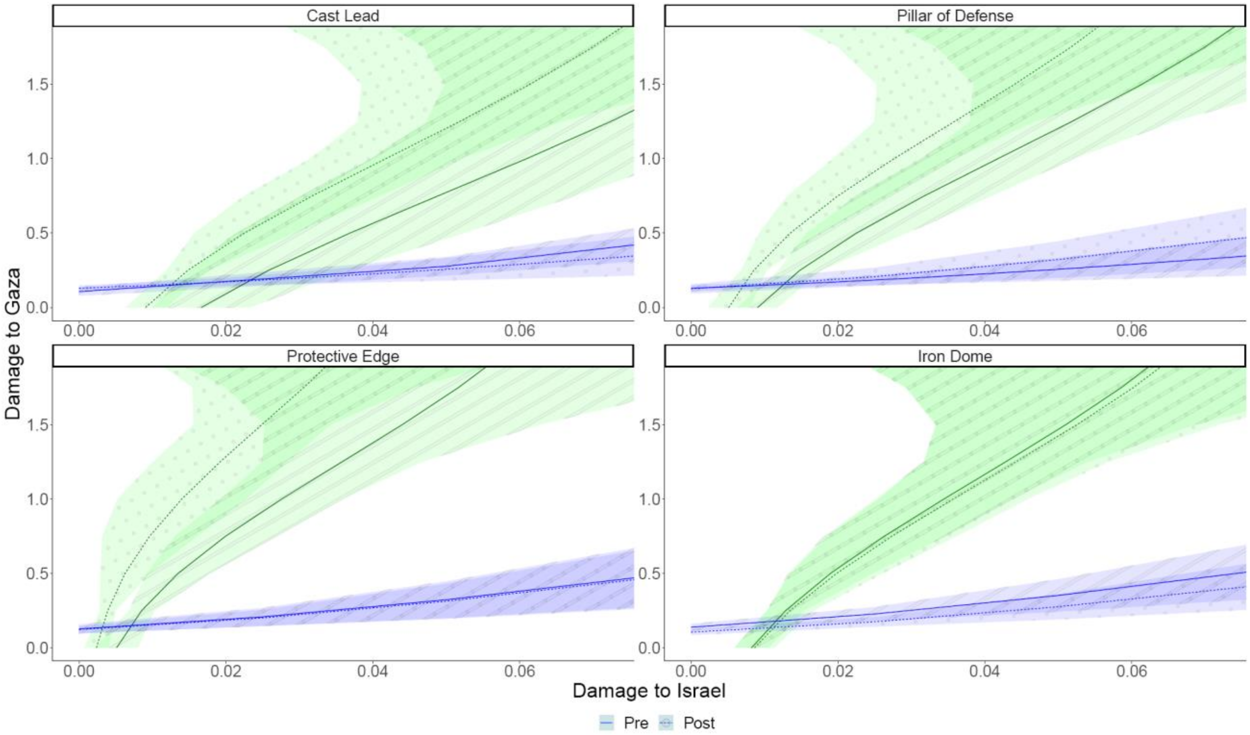
7. Major Operations and Iron Dome Missile Defense

We turn now to interventions that could shift response curves, the three operations in our sample period characterized by periods of high frequency attacks initiated by Israel, accompanied by ground incursions in Gaza, as well as the introduction of the Iron Dome anti-missile defense system. These are not escalations (though they are often mischaracterized as such), since they are not the result of a sequence of escalating actions (as in Figure 2) but instead pre-planned offensive operations that begin with a large violent action. During these periods the objective of the IDF, in its description, is to destroy some capacity of the other side to use violence --to deter by denial, as we discussed in Section 2B above.

As illustrated in Figure 4, in our response curve approach, deterrence by denial shifts an opponent's response curve to a less aggressive stance (i.e., a smaller intercept and perhaps less slope) thereby achieving a less violent equilibrium. So, revisiting the fifth hypothesis of Section 6A, we can test whether major operations achieve deterrence by denial by seeing if the response curve of Gazan militants shifts to the left.

Figure 9 illustrates that test, implemented for each of the major operations. In each figure we report estimates of both Gazan and Israeli response curves allowing for a shift between the period before and after those operations –recalling that data from the period of the operations is excluded from our sample. The shift is illustrated by movement from the solid to the serrated line, for Cast Lead (November 2008 – January 2009, upper left), Pillar of Defense (November 2012, upper right), and Protective Edge (June 2014 – August 2014, lower left). In all three cases the Gazan response curve shifts to lower damage caused at all levels of damage suffered. The standard error bands imply that in the case of Pillar of Defense and Protective Edge that shift is statistically significant in the neighborhood of the pre-operation equilibrium. The Israeli curve does not display a significant shift. In all three cases the new equilibrium is less violent, shifting to significantly lower violence suffered by Israel, and lower violence suffered by Gazans (though not statistically significantly so).

Figure 9: Do Major Operations Achieve Deterrence by Denial?



Finally, the bottom right panel reports on the same exercise for the Iron Dome missile defense system. Our reasoning in Section 2B above was that the Iron Dome system would increase the marginal cost to Gazan militants of imposing damage on Israel at all levels of damage, leading to a shift of the Gazan response curve to less violence (i.e., to the left). Instead, there seems to have been no shift at all. Shifting that equilibrium was not the only purpose of the very expensive Iron Dome system, which also provides psychological and physical security to people, hospitals and other critical institutions. Yet the result is still surprising. We leave analysis of it to future research.

8. Conclusions

This paper studies deterrence in the context of the Israel-Gaza conflict, estimating response curves in a repeated game. Estimation is enabled by a uniquely detailed dataset coding daily reports compiled by United Nation peacekeepers covering the period 2007 through 2017. The conflict is characterized by frequent, short episodes of violence. We code 1,168 episodes, of which the median lasts less than 5 hours, but the median pause between episodes is less than 5 days. An episode averages about 6 actions by each side. On average 0.13 Gazans are killed per episode and 0.52 injured, while the figures for Israelis are much lower: 0.004 and 0.04.

Our analysis reveals that this low level of persistent violence can be understood as the equilibrium of a repeated game with mixed strategies, which we solve by estimating expected response curves for each side and solving their crossing point. Estimates imply that the crossing point is unique and stable: sequences of attacks and counterattacks tend to converge to it. Yet that equilibrium is stochastic, pulling sequences toward it in probability but not holding them at a single point.

Both response curves are upward-sloping, indicating a deterrent posture. Yet the equilibrium achieved displays *incomplete deterrence*, in the sense that the stable equilibrium remains violent (i.e., is not at the origin). This characterization is novel and relevant, --in our view, to many other security situations in which deterrence is incomplete: cyber-warfare, espionage, disinformation campaigns, and other grey zone conflict.

Major operations by Israel are shown to achieve *deterrence by denial*, reducing the violence of the resulting equilibrium by shifting the Gazan response curve to less violence.

Though the data reveal some evidence that episodes are not fully independent, the repeated game paradigm provides an excellent approximation of these data, describing the conflict as the combatants do. That description contrasts with the empirical literature on this conflict which may overstate the duration (and damage) due to sequences of actions set off by an individual attack. Those studies were set in a previous period, and included the West Bank, which may explain the contrast, yet we suspect that they reflect a bias of the VAR approach.

We believe that our major innovation, estimating response curves within episodes of conflict, can illuminate analysis of potential escalation and de-escalation in conventional conflicts. It also holds potential for policy-relevant studies of deterrence in “gray zone” fields, such as cybersecurity and disinformation, where frequent, repeated interaction is common.

In future work we plan to proceed in multiple directions: using observed responses to better understand how sides weigh civilian casualties as opposed to casualties among their forces; and analyzing “frenemies” who simultaneously engage in positive-sum exchanges (e.g., trade and climate agreements) and negative-sum activity, such as attacks, cyberwarfare, and espionage.

REFERENCES

Abrahams, Alexei and Berman, Eli and Khadka, Prabin and Klor, Esteban F. and Powell, John. 2019. "Mostly Deterred: An Episodic Analysis of The Israel-Gaza Conflict" (October). http://dx.doi.org/10.2139/ssrn.3465438

Asali, Muhammad, Aamer S. Abu-Qarn, and Michael Beenstock. 2017. "The Cycle of Violence in the Second Intifada: Causality in Nonlinear Vector Autoregressive Models," *Journal of Applied Econometrics*, 32 (6): 1197–1205.

Axelrod, Robert. 1984. *The Evolution of Cooperation.* New York: Basic Books.

Baliga, Sandeep and Tomas Sjöström. 2009. "Contracting with Third Parties," *American Economic Journal: Microeconomics*, 1(1): 75-100.

Bajari, Patrick, C. Lanier Benkard, and Jonathan Levin. 2007. "Estimating Dynamic Models of Imperfect Competition." *Econometrica* 75 (5): 1331–70. https://doi.org/10.1111/j.1468-0262.2007.00796.x.

Benmelech, Efraim, Claude Berrebi and Esteban F. Klor. 2015. "Counter-Suicide-Terrorism: Evidence from House Demolitions," *Journal of Politics*, 77 (1): 27-43.

Bergman, Ronen and Adam Goldman, "Israel Knew Hamas's Attack Plan More than a Year Ago," *New York Times*, December 2, 2023, https://www.nytimes.com/2023/11/30/world/middleeast/israel-hamas-attack-intelligence.html. (Downloaded January 17, 2024.)

Berman, Eli and David D. Laitin. 2008. "Religion, Terrorism and Public Goods: Testing the Club Model," *Journal of Public Economics*, 92 (10–11): 1942–1967.

Berman, Eli, David A. Lake, Gerard Padró I Miquel and Pierre Yared. 2019. "Principals, Agents, and Indirect Foreign Policies," in *Proxy Wars: Suppressing Violence Through Local Agents*, Eli Berman and David A. Lake (eds.). Ithaca, NY: Cornell University Press.

Chassang Sylvain and Gerard Padró I Miquel. 2010. "Conflict and Deterrence Under Strategic Risk," *The Quarterly Journal of Economics*, 125 (4): 1821-1858.

Dugan, Laura and Erica Chenoweth. 2012. "Moving Beyond Deterrence: The Effectiveness of Raising the Expected Utility of Abstaining from Terrorism in Israel," *American Sociological Review*, 77 (4): 597-624.

Durante, Ruben and Ekaterina Zhuravskaya. 2018. "Attack When the World is Not Watching? International Media and the Israeli-Palestinian Conflict," *Journal of Political Economy*, 126 (3): 1085-1133.

Etkes, Haggay and Assaf Zimring. 2015. "When Trade Stops: Lessons from the Gaza Blockade 2007-2010," *Journal of International Economics*, 95 (1): 16-27.

Evron, Yair. 2005. "Disengagement and Israeli Deterrence," *Strategic Assessment*, 8 (2): 6-11.

Fearon, James D. 1994. "Domestic Political Audiences and the Escalation of International Disputes," *The American Political Science Review*, Vol. 88 (3): 577–592.

Fearon, James D. 1997. "Signaling Foreign Policy Interests Tying Hands versus Sinking Costs," *Journal of Conflict Resolution*, 41(1): 68–90.

Fearon, James D. 2002. "Selection Effects and Deterrence," *International Interactions*, 28(1): 5–29.

Freedman, Lawrence. 2004. *Deterrence*. Cambridge, UK: Polity.

Fudenberg, Drew and Jean Tirole. 1991. *Game Theory*. Cambridge MA: The MIT Press.

Getmansky, Anna and Thomas Zeitzoff. 2014. "Terrorism and Voting: The Effect of Rocket Threat on Voting in Israeli Elections," *American Political Science Review*, 108 (3): 588-604.

Granger, Clive W. J. 1969. “Investigating Causal Relations by Econometric Models and Cross-Spectral Methods,” *Econometrica*, 37 (3): 424-438.

Gibilisco, Michael. 2023. “Mowing the Grass,” *Journal of Theoretical Politics*, 35 (3): 204-231.

Golan, David and Jonathan D. Rosenblatt. 2011. “Revisiting the Statistical Analysis of the Israeli–Palestinian Conflict,” *Proceedings of the National Academy of Sciences*, 108 (15).

Heise David. 1989. Modeling Event Structures. *Journal of Mathematical Sociology.* 14:139–69.

Hubbard, Ben and Maria Abi-Habib, “Behind Hamas’s Bloody Gambit to Create a ‘Permanent’ State of War,” *New York Times*, Nov. 8, 2023

Haushofer, Johannes, Anat Biletzki and Nancy Kanwisher. 2010. “Both Sides Retaliate in the Israeli-Palestinian Conflict,” *Proceedings of the National Academy of Sciences*, 107 (42): 17927-17932.

Hila’I, Sivan, 2023. Netanyahu in 2017: “Hamas Planned an Attack with Missiles, Hang-gliders, Kidnapping and Murder in Civilian Settlements” Ynet, December 12, 2023. https://www.ynet.co.il/news/article/ryd2lfgpp (downloaded January 17, 2024).

Inbar, Efraim and Eitan Shamir. 2014. “Mowing the Grass: Israel’s Strategy for Protracted Intractable Conflict,” *Journal of Strategic Studies*, 37 (1): 65-90.

Jaeger, David A., Esteban F. Klor, Sami H. Miaari and M. Daniele Paserman. 2012. “The Struggle for Palestinians Hearts and Minds: Violence and Public Opinion in the Second Intifada,” *Journal of Public Economics*, 96 (3-4): 354-368.

Jaeger, David A., Esteban F. Klor, Sami H. Miaari and M. Daniele Paserman. 2015. “Can Militants Use Violence to Win Public Support? Evidence from the Second Intifada,” *Journal of Conflict Resolution*, 59 (3): 528-549.

Jaeger, David A. and M. Daniele Paserman. 2006. “Israel, the Palestinian Factions and the Cycle of Violence,” *The American Economic Review*, 96 (2): 45-49.

Jaeger, David A. and M. Daniele Paserman. 2008. “The Cycle of Violence? An Empirical Analysis of Fatalities in the Palestinian-Israeli Conflict,” *The American Economic Review*, 98 (4): 1591-1604.

Jaeger, David A. and M. Daniele Paserman. 2009. “The Shape of Things to Come? On the Dynamics of Suicide Attacks and Targeted Killings,” *The Quarterly Journal of Political Science*, 4 (4): 315-342.

Jervis, Robert, Richard N. Lebow and Janice G. Stein. 1985. *Psychology and Deterrence*. Baltimore, MD: Johns Hopkins University Press.

Kahn, H. 1960. *The Nature and Feasibility of War and Deterrence*. Rand Corporation.

Kydd, Andrew and Barbara F. Walter. 2002. “Sabotaging the Peace: The Politics of Extremist Violence,” *International Organization* 56(2): 263–296.

Lichbach, Mark Irving, 1987. “Deterrence or Escalation? The Puzzle of Aggregate Studies of Repression and Dissent,” *The Journal of Conflict Resolution*, Vol. 31, No. 2 (June), pp. 266-297.

Leventoğlu, Bahar, and Branislav L. Slantchev. 2007. “The Armed Peace: a Punctuated Equilibrium Theory of War,” *American Journal of Political Science*, 51 (4): 755–771.

MacKinnon, James G., Morten Ørregaard Nielsen, and Matthew D. Webb. 2023. Testing for the Appropriate Level of Clustering in Linear Regression Models. *Journal of Econometrics* 235 (2): 2027–56.https://doi.org/10.1016/j.jeconom.2023.03.005.

Marlin-Bennett, Renee, Alan Rosenblatt, and Jianxin Wang. 1991. "The Visible Hand: The United States, Japan, and the Management of Trade Disputes." *International Interactions* 17:191-213.

Mearsheimer, John J. 1983. *Conventional Deterrence*. Ithaca, NY: Cornell University Press.

Moore, Will H. 1998. “Repression and Dissent: Substitution, Context, and Timing,” *American Journal of Political Science*, Vol. 42, No. 3 (July), pp. 851-873

Myerson, Roger. 2008. “Leadership, Trust, and Power: Dynamic Moral Hazard in High Office.” Working Paper.

Nanes, Matthew J. 2019. “Lebanon and Gaza, 1975-2017: Israel’s Extremes of Interest Alignment,” in *Proxy Wars: Suppressing Violence Through Local Agents*, Eli Berman and David A. Lake (eds.). Ithaca, NY: Cornell University Press.

Padró I Miquel, Gerard and Pierre Yared. 2012. “The Political Economy of Indirect Control,” *The Quarterly Journal of Economics*, 127 (2): 947-1015.

Powell, R. 1990. *Nuclear Deterrence Theory, the Search for Credibility*. Cambridge MA: Harvard University Press.

Powell, Robert. 1999. *In the Shadow of Power: States and Strategies in International Politics.* Princeton. Princeton, NJ: Princeton University Press.

Powell, Robert. 2003. “Nuclear Deterrence Theory, Nuclear Proliferation, and National Missile Defense,” *International Security*, 27(4): 86–118.

Powell, Robert. 2012. “Persistent Fighting and Shifting Power,” *American Journal of Political Science*, 56(3): 620–637.

Rubin, Uzi. 2011. “The Missile Threat from Gaza: From Nuisance to Strategic Threat,” *The Begin – Sadat Center for Strategic Studies*, Bar Ilan University, Mideast Security and Policy Studies No. 91.

Schelling, Thomas C. 1960. *The Strategy of Conflict*. Cambridge, MA: Harvard University Press,

Schelling, Thomas C. 1966. *Arms and Influence*. New Haven, CT: Yale University Press.

Sobelman, Daniel. 2017. “Learning to Deter: Deterrence Failure and Success in the Israel-Hizbollah Conflict, 2006-2016,” *International Security*, 41 (3): 151-196.

United Nations Office for the Coordination of Humanitarian Affairs (OCHA). 2016. *The Gaza Strip: The Humanitarian Impact of the Blockade*, November.

Wooldridge, Jeffrey M., 2010. *Econometric Analysis of Cross Section and Panel Data* (Second Edition), Cambridge MA: MIT Press.