The Effect of Democratic Accountability on Counter-Terrorism

Olga Chyzh*

Abstract

In the aftermath of major security incidents, the public often supports retaliation against the perpetrators and additional security—measures that do not always reduce the threat of future security incidents. Security measures, moreover, are often dual-purpose: they may be used against legitimate security threats or domestic dissent. Under what conditions does democratic accountability come in conflict with public safety? And can the public constrain a predatory incumbent from misrepresenting the level of threat to enhance their chances of survival in office? I answer these questions by analyzing the interaction between a democratic incumbent, the public, and an extremist group as a three-player strategic game. The model reveals two scenarios, in which public accountability may enable outcomes that are suboptimal from the perspective of public safety. In one, a non-predatory government overreacts against a policy-oriented extremist group, risking escalation; in the other, a predatory leader manipulates threats to suppress democratic freedoms. Both cases show that public demands for retaliation may unintentionally drive governments toward inefficient security policies or even to undermining democratic checks. I illustrate the theoretical logic and insights using the case of Turkey between 1984–2012.

*Associate Professor of Political Science, University of Toronto, http://www.olgachyzh.com/

The October 7, 2023, terrorist attacks on Israel served as a stark reminder that terrorism remains a persistent threat to global and national security. The discourse surrounding these events, however, has markedly diverged from the typical mixture of outrage on behalf of victims and calls for retaliation against perpetrators. Instead, much of the discussion—including that by Israeli leader Netanyahu's domestic audience—has centered on incisive critiques of the Israeli government's policies and motives. Critics have gone as far as to charge Netanyahu with using the terrorist threat and the ensuing invasion of Gaza to extend his stay in power (New York Times 2024; The Atlantic 2024; Allon 2024). These events, and the surrounding discourse highlight a key tension between a democratic leader's private goal of maximizing electoral prospects and the public mandate to defend the country from extremist attacks.

How did a democracy like Israel end up in such a situation? Why did democratic institutions fail to prevent this crisis? And more broadly, under what conditions can a democratic public constrain an incumbent who may be using an extremist threat to stay in power? I answer these questions by developing a three-player game among a democratic government, the public, and a domestic extremist group.

In the game, the government's central function is to solve the guns-vs-butter dilemma, which consists of optimizing the allocation of state resources towards two goals: public safety and public welfare. The dilemma arises as a result of a disconnect between what this optimum allocation looks like from the perspective of the public vs. the regime. The public prefers that the government distribute resources based on an honest assessment of the existing security threat level: that is, prioritize defense in the presence of a security threat, and prioritize welfare otherwise. The regime, in contrast, may have a private preference for prioritizing defense over welfare, as many defense measures (e.g., intelligence gathering, restrictions on movement or speech) are easily re-purposed from containing legitimate security threats to constraining domestic opposition. As long as the public has a mechanism to replace the

leader, regime survival is predicated on its ability to balance between these two competing sets of preferences.

The government's biggest advantage is informational: the public is not privy to the intelligence that would reveal the real level of security threat. Given it's private preference for implementing defense measures, the government has an incentive to misrepresent or exaggerate the security threat, so as to justify additional defense measures. In the wake of the Maidan protests in 2013–2014, for example, Ukrainian President Yanukovich was quick to label the protesters as "terrorists" to justify shutting down the subway in central Kyiv (CBC 2014). The danger, of course, is that always exaggerating the threat risks losing the public's trust. If the government's claim of a security threat is not credible to the public, then a better survival strategy may be focusing on policy provision.

Despite the government's assertions, not all domestic extremist groups are equally threatening: some groups pursue maximalist goals, whereas others could be satisfied with limited policy concessions. A group's level of threat, moreover, may be endogenous to government policy. Responding to an extremist group's policy demands with militarized actions may lead to escalation that could have been avoided by a conciliatory strategy. Some groups, including the African National Congress (ANC), Irish Republican Army (IRA), and the Kurdistan Workers' Party (PKK), have agreed to and upheld ceasefires during negotiations with the government or in the aftermath of peace settlements.

The game's major innovation and a departure from the previous literature is that the game treats the public as a strategic actor whose payoff parameters go beyond the basic needs of safety and well-being. In stride with recent literature that highlights emotion as a key lens through which the public views terrorism, the public's payoff function is specified to include a value for retaliation against extremist groups who commit attacks (Huddy et al. 2005; Wayne 2023; Schnakenberg and Wayne 2024). In addition, the public's payoffs also include a preference for good governance, that is a leader driven by the goal of public welfare

rather than their own private benefit from staying in power.

The game's central finding is that the public's plight is often self-inflicted. All leaders want to stay in power, autocrats and democrats alike, and democratic checks and balances are not foolproof insurance of good governance or nonpredatory leaders. In fact, democratic accountability may, under some conditions, induce leaders to implement predatory or inefficient policy. When the public clamors for retaliation, any electorally-motivated government has a strong incentive to implement additional, even inefficient security measures, irrespective of whether the government itself is predatory.

Two scenarios emerge. In one, a non-predatory government implements unnecessary and counter-productive security measures against a policy-oriented group that would have accepted concessions. In the other, a predatory government does so, while also exploiting this as an opportunity to curtail existing democratic constraints. In the former scenario, the costs of public's impulsiveness is a short-term increase in the risk of extremist attacks. In the second scenario, the costs are compounded by long-term damage to the country's democratic institutions. To make matters worse, these conditions may also lead to the incumbent losing the support of the pivotal voter, which may force a choice between removal from power or democratic backsliding.

In either case, the game shows, counter-intuitively, that the public's preference for retaliation is a sufficient condition, under which even a non-predatory government pursues a militant strategy against policy-oriented extremists, who would have accepted policy concessions. This finding is grounded in the inherent tension between public safety and public opinion: evidence shows that the public tends to support security policies, such as increase in defense spending or restrictions on civil liberties and privacy, in the wake of major security incidents (Merolla and Zechmeister 2009; Huddy and Feldman 2011; Malhotra and Popp 2012; Brooks and Manza 2013) or when such policies are motivated by the goal of counter-terrorism (Conrad et al. 2018). Ironically, these measures do not always increase

future security from terrorist attacks (Dragu 2011).

The model elucidates the causal processes that may lead to failure of democratic constraints on the elected leaders, such as the one described in the opening example. Though the model yields a plethora of novel insights, the multiplicity of equilibria precludes direct empirical tests. Rather than hypothesis testing, I use empirical data to illustrate the game's logic and results. The primary case studies include four periods of interactions between the Turkish government and two extremist groups: the PKK and Ergenekon.

The article proceeds in the following way. After situating this research within the recent literature, I present the model, discuss its insights, and describe the equilibria using a temporal case study of Turkey, 1984–2012. I conclude by discussing future directions and policy implications.

Democracy and Political Extremism

There is an inherent contradiction between a democratic government and domestic political extremism. On one hand, a democratic government pre-supposes societal group's ability to act on their political demands via legitimate channels of political representation. The formation of a domestic extremist group, on the other hand, implies a failure of the democratic process, a denial of political rights to a societal group. An extremist group's survival and endurance further implies support from a sizeable portion of the public (O'Connor 2021), In a democratic context, this means that political extremists draw their support from a portion of the electorate. Therefore, in a democracy, an extremist group may influence the outcome of an election even in the absence of electoral prospects of their own—by chipping away votes from the viable political candidates, especially the incumbent. Furthermore, portions of the electorate may turn against the incumbent, not out of their support for the extremists, but due to their disapproval of the incumbent's handling of the issue, e.g., they may

prefer that the government address the extremists' grievances through policy rather than counter-insurgent measures.

Political violence research has focused on the competing incentives faced by the government and opposition groups, as well as on the role of domestic public. In some research, the government and the opposition is modeled as competing for the support of the public. In these studies, repression does not always deter political dissent (Ritter 2014; Ritter and Conrad 2016) and may even push more sympathizers to join or support the dissent (Bueno de Mesquita and Dickson 2007). In the terrorism literature, there is a similar debate on the effectiveness of a militant counter-terrorism response to terrorist attacks (Schneider, Brück, and Meierrieks 2015). While Lyall (2009) contends that indiscriminate shelling decreases future attacks, others argue that a heavy-handed counterterrorism response may drive additional supporters into the terrorist ranks (Kydd and Walter 2006; Findley and Young 2012).

A different approach is to model government response to the threat of terrorism as a signaling game, in which the government decides whether to concede or resist by inferring the terrorists' capabilities from the history of attacks (Lapan and Sandler 1993; Overgaard 1994; Pape 2003). The literature disagrees on the effectiveness of terrorist attacks on eliciting policy concessions. Some argue that the use of terrorist tactics reduces the probability of negotiated settlements (Fortna 2015) or concessions (Abrahms 2006). Thomas (2014), in contrast, argues that, beyond conveying the terrorists' strength, successful attacks essentially put the state in a tough position where negotiations and concessions are but the only option.

Concessions may entice additional disgruntled groups to organize and issue their demands (Walter 2006*a*,*b*). Bueno De Mesquita (2005) models the government's offers to different factions within terrorist organizations, such as the moderates vs. the extremists, to show how concessions may radicalize the group by leaving it in the hands of the extremists. Cunningham (2011) argues that governments may use concessions as a strategy to divide and weaken a group by appeasing the moderates and isolating the extremists.

Other work has focused on the tension between increased security and the public safety from the threat of terrorism. Dragu and Polborn (2014), for example, show that the electoral incentives may entice an unconstrained executive to implement more aggressive counter-terrorist policies than are necessary or effective, given the actual terrorist threat. Dragu (2011, 2017) also show that reduction in privacy protection, such as collection of information on private citizens, may in fact decrease, rather than increase, the domestic terrorist threat. Others link counterterrorism to the party of the incumbent, demonstrating that left-wing governments may respond to the threat of terror more aggressively than right-wing governments (Berrebi and Klor 2006; Di Lonardo 2017).

These studies, however, leave out a key theoretical tension related to the role of the democratic public in retaining the incumbent. Whenever the public is treated as a strategic actor, its payoff function is usually reduced to the basic needs of safety and well-being. In actuality, a democratic public's support for the government's chosen counter-terrorism strategy depends on two other important parameters—the public's value of retaliation (Wayne 2023; Schnakenberg and Wayne 2024), and the strength of the anti-corruption sentiment.

First, extremist attacks evoke emotions of anger in the members of the public, who often react by calling for retaliation against the perceived attacker (Fisk, Merolla, and Ramos 2019; Wayne 2023). The goal of retaliation, however, may conflict with that of reducing future attacks, by setting off a cycle of violence between the state and the extremists. Second, the public assesses specific government policies, such as effectiveness of their counter-terrorism strategies within the broader context of government performance. Public perception of corruption decreases government support (Wang and Dickson 2022). If the public perceives the government as corrupt and ineffective, it may hold new counter-terrorism measures, such as expansion of police powers or surveillance, to greater scrutiny than when the public trust in government is high (Denemark 2012). The model fills this gap by incorporating these important parameters into the public's payoff function and exploring their implications.

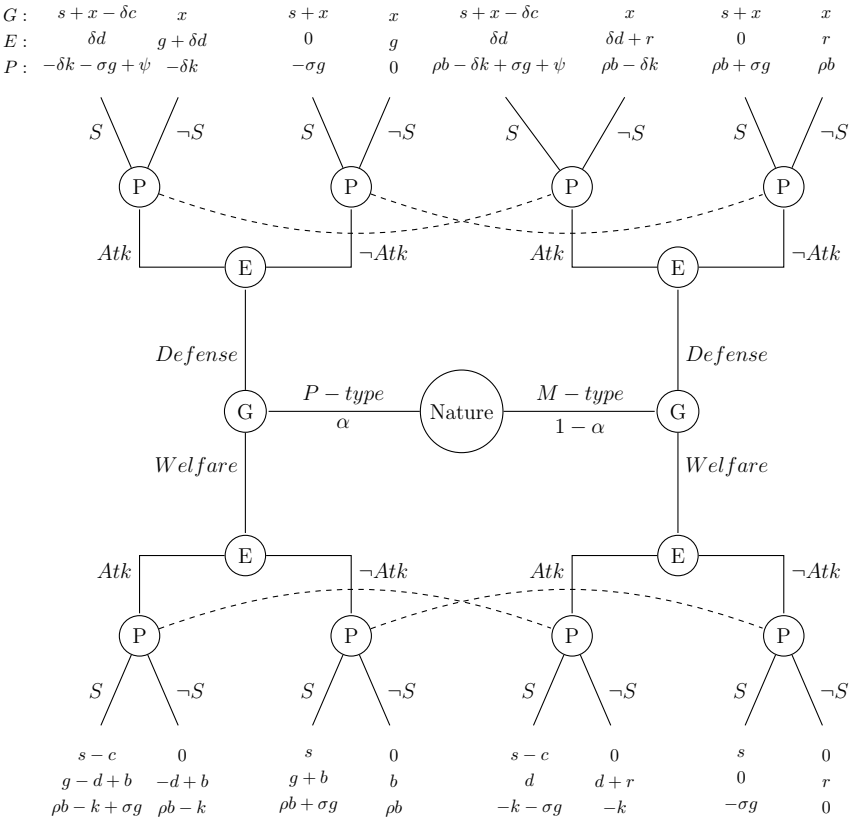
A Theoretical Model of Counterterrorism

The game focuses on the interaction between three actors: *Government* (**G**), *the Public* (**P**), and a radical anti-government group, *Extremists* (**E**). The government is the head of the executive office in a country. The public is a sizeable subset of the domestic population, whose support is necessary for the government to stay in power. Alternatively, one may think of the public as the pivotal voter whose support the leader must win in order to stay in power through a democratic process. I use these two terms inter-changeably. The *Extremists* are a group of individuals within the society that is willing to pursue its goals through premeditated use of violence.

In line with previous game-theoretic research, the game features an informational asymmetry among the players regarding the type of the extremists (Arce and Sandler 2007, 2010). Specifically, the extremists and the government are more informed than the public. While the government is able to correctly assess the type of the extremists, the public has incomplete information about whether the domestic opposition group is merely seeking concessions on a set of policies (*p-type*) or are an uncompromising militant group with extreme demands (*m-type*). *P-type* groups seek policy concessions from the government, are open to negotiations, and generally use violence only as a last resort. In contrast, *m-types* are a fanatical group that derives an added utility from the use of violence itself (Arce and Sandler 2007, 2010). According to these definitions, examples of *p-type* extremists may include such groups as Euskadi ta Askatasuna (ETA), ANC, and the Tamil Tigers, while *m-types* extremists may include al Qaida, the Islamic State (ISIS), or the Abu Nidal Organization (ANO). The primary examples of *p-type* extremists discussed in this paper are the PKK in Turkey, the IRA in Northern Ireland, and the PLO (Palestinian Liberation Organization).

Figure 1 depicts the game in extensive form, along with each actors' payoffs. The game starts with *Nature* (N), determining the type of the domestic extremists as either a polit-

Figure 1: An Extensive-Form Game between the Government, the Public, and the Terrorists



ically motivated *p-type* with probability α, or a militant *m-type* with probability $(1-\alpha)$. The extremists know their own type, and so does the government, while the public is only informed about the mean of the distribution of the types in the population, α.

The Government's Choice

The government knows the type of the extremists and has to choose between implementing a *Defense* (D) or a *Welfare* (W) policy. This decision involves the classic "guns-vs-butter" trade-off, i.e. a greater focus on defense takes resources away from policy provision. From the perspective of the government, this decision also has a second, private trade-off. Expanding defense measures increases the government's information on the activities of the domestic opposition and provides tools for preempting dissent. In the context of the model, policy

provision refers to the types of policies that would benefit both the public and the supporters of the extremist group, albeit to different degrees. These policies may include schools, language protections, or other minority rights.

The government prefers to stay in power and, for this, values the support of the public. To reflect this incentive, the government's benefit from the support of the public is modeled via a positive parameter, s, which appears in the government's payoffs from all outcomes in which the public supports the government. The government also derives a benefit from choosing *Defense* over *Welfare*, as modeled via a positive parameter, x, that accrues to the government each time it chooses *Defense* over *Welfare*. In case of an extremists' attack, the government pays a cost, c, which is reduced to δc, $(0 < \delta < 1)$, if the government chose *Defense*. That is, implementing counter-terrorist policies reduces the damage the extremists can inflict.

Given this incentive structure, the government's ideal outcome in this game is to convince the public that the extremists are of the *m-type*, so that it can justify the *Defense* policy while preserving the support of the public.

The Extremists' Choice

After the *Government* moves, the next move is by the extremists. The extremists know their own type, and can observe the *Government*'s action. The extremists' choose between *Attack* (Atk) and *No Attack* ($\neg$Atk), but the two types face different incentives and payoff structures.

P-type Extremists *P-type* groups align with a large segment of the public in terms of their policy goals. Unlike the public, they are willing to resort to violent means to achieve these goals or punish the government for acting in a predatory manner. This is modeled via a b benefit that enters into *p-type*'s payoffs for all outcomes in which the government

implements a welfare policy. As long as the government implements welfare policies, *p-type* groups do not benefit from engaging in violence and prefer that the public support the government (i.e., that the public play S). This is reflected in the $-d$ parameter in *p-type*'s payoff from any outcome in which they engage in violence against a government that implements welfare policies and in the g parameter in *p-type*'s payoff from any outcome in which the public supports a nonpredatory government.

If the government implements a *Defense* policy against a *p-type* group, then the group derives a benefit from punishing the government by engaging in violence, as modeled via parameter δd that enters into *p-type*'s payoff from all outcomes in which they attack a predatory government. The benefit from attacking, d, is reduced by the amount δ $(0 \geq \delta \geq 1)$ as the government's implementation of a *Defense* policy helps reduce the damage from an attack. *P-types* also derive an additional benefit if the public removes the predatory government from office, as is reflected by the g parameter that enters into the *p-type*'s payoffs from every outcome in which the public does not support a predatory government.

M-type Extremists *M-type* extremists are a militant group whose primary goal is to engage in violence against the government and the public, i.e. *m-types* always prefer attacking to not attacking. If the government chooses *Welfare*, then *m-type*'s payoff from attacking contains a benefit d, while if the government chooses *Defense*, *m-type*'s payoff from attacking decreases to δd, $(0 < \delta < 1)$. If the public does not support the government, then *m-type*'s utility also contains a benefit r, which can be thought of as the benefit associated with additional political instability, or a possible increase in the extremists' funding and/or recruitment.

The Public

The last move is by the *Public*. The public is uninformed about the type of the *Extremists*, but observes the actions of the *Government* and the *Extremists*. The *Public*'s choice is whether to *Support* the *Government* (S) or *Not Support* the *Government* ($\neg$S).

The public prefers that the government choose *Defense* when the extremists are of the *m-type* and *Welfare* otherwise, which is modeled via a benefit ρb (policy benefit). The public's policy benefit ρb is positively correlated with the policy benefit obtained by the *p-type* extremists. That is, while the public's preferences are not in perfect alignment with those of the extremists, the *p-type* extremists are advocating for the policies that would benefit the public, albeit using the violent means that the public does not condone.

The public obtains a benefit σg for supporting a non-predatory/competent government[1], and a cost $-\sigma g$ for supporting a predatory/incompetent government. Again, the public's "good governance" benefit σg is positively correlated with the *p-type* extremists' "good governance" benefit, albeit the two benefits need not be the same.

In case of an attack, the public also pays a cost of k (if the government chose *Welfare*) or δk, $(0 < \delta < 1)$ (if the government chose *Defense*). The public also derives a "retribution" benefit, ψ, as long as they supported the government and the government counter-terrorism measures were followed by an attack.

The public's task is to correctly guess the type of the extremists from the government's signal and punish the government if they are acting in a predatory or incompetent manner. The caveat is that, when the government acts in a predatory manner, *p-type* extremists may act indistinguishably from *m-type* extremists, which complicates the public's decision.

The difference in preferences between the public and *p-types* is that, while both prefer to remove a predatory government, the public prefers the outcome in which the government

[1]A predatory government chooses *Defense* when the terrorists are of the *p-type*; an incompetent government chooses *Welfare* when the terrorists are of the *m-type*.

Table 1: Game Parameters

Parameter	**Description**	**Constraints**
α	Probability that the extremists are of the *p-type*	$0 < \alpha < 1$
δ	Added security resulting from *Defense*	$0 < \delta < 1$
	Government's Payoffs:	
c	Cost of a terrorist attack	$c > 0$
s	Benefit from gaining support of the pivotal voter	$s > 0$
x	Added security in office derived from *Defense*	$x > 0$
	Extremists' Payoffs:	
b	Policy benefit	$b > 0$
d	Benefit from attacking the government	$d > 0$
g	Good governance benefit	$g > 0$
r	"Chaos" benefit from removing a non-predatory government	$r \geq 0$
	Public's Payoffs:	
k	Cost of a terrorist attack	$k > 0$
ρb	Policy benefit	$\rho \geq 0$ $b > 0$
σg	Good governance benefit	$g > 0$,$\sigma > 0$
ψ	Retribution benefit for fighting terrorism	$\psi \geq 0$
Additional Constraints		
$d > g > \delta d$		
$d > \delta d > r > \delta r$		
$\sigma g > \psi$		
$s > x > \delta c$		

is removed peacefully, while the *p-types* prefer to also punish a predatory government by engaging in a violent attack against it. That is, from the perspective of the public, a violent removal carries additional costs, such as destruction of property, loss of life, and political and economic uncertainty.

Table 1 provides a list of all game parameters and constraints. Most of the constraints purely ensure that the results are within reasonable values, e.g., that probabilities are constrained within the interval of $[0, 1]$. The only nontrivial constraint is $s > x > \delta c$, so as to prevent the government's *Welfare* strategy from being strictly dominated by *Defense*.

Equilibria Analysis

The Appendix contains the full solution, using the Perfect Bayesian Equilibrium solution concept. The game has three equilibria classes: separating, pooling, and semi-separating. Figure 2 shows the observed (solid lines) and latent probabilities of the government choosing D in response to *p-types* (black), *p-type* extremists attacking in response to D (green), the public supporting the government after observing $\{D, Atk\}$ (blue), and the effect of these actions on government loss of public support (red) for a set of plausible parameter values, while varying the public's concern with government corruption, g.[2] The subfigure on the left shows the equilibria in which the government plays pure strategies (EQ1, EQ2, EQ3, and EQ5); the subfigure on the right shows the equilibrium in which the government mixes its strategies against the *p-types* (EQ4), in which the government plays a mixed strategy against *p-type* extremists.

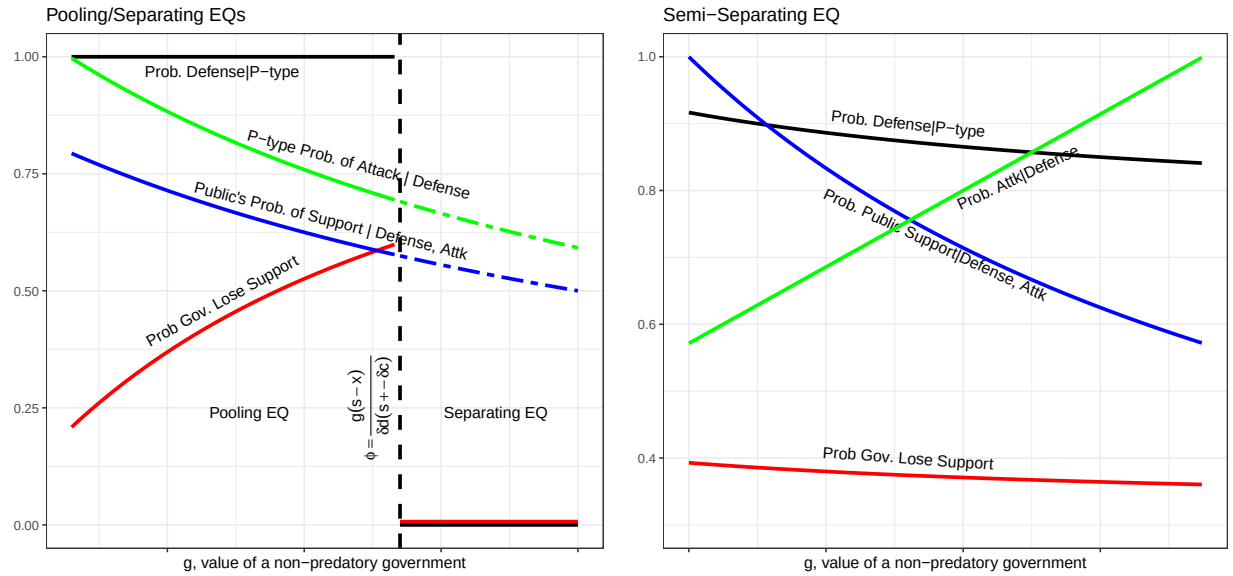
In the first subfigure, the black vertical dashed line denotes the threshold value of g that separates the support for the pooling vs separating equilibria: pooling on the left, separating on the right. In the pooling equilibria, the government prioritizes its survival in office over its mandate to protect the public. No matter the extremist type, the government always plays D, despite the risk of attacks by *p-type* extremists and even losing the support of the public.[3] The pooling equilibria holds for rather high probability of attack by the *p-type* extremists (as high as 1).[4] Also note-worthy, the pooling equilibrium exists even when the public only supports the government after a game history of *Defense, Attack* with a rather high probability. In other words, the government is willing to act on its predatory incentive even with high chance of losing the support of the public.

[2] *M-type* extremists always play Atk, and the government always plays D so these probabilities are not shown.

[3] Note that in the game, the support of the public is a sufficient but not necessary condition for survival. That is, an incumbent can also use x to survive by undermining democratic checks on their power.

[4] The analysis focuses on the strategy for the *p-type* extremists, as the *m-type*' payoff structure makes not attacking always a strictly dominated strategy.

Figure 2: The Effect of the Retaliation Parameter on Players' Strategies and Government Support



Note. Black line shows the probability that the government plays D against *p-types*, green denotes the probability that the *p-type* extremists attack in response to defense, blue is the probability that the public support the government after a game history of $\{D, Atk\}$, and red the probability of a game outcome in which the public does not support the government. Dashed lines indicate unobservable parts of the equilibria, solid lines indicate observable parts. All parameters are set at reasonable fixed values.

In contrast, the separating equilibria are the "good governance" equilibria: the government ignores the predatory incentive and fulfills its mandate to act in the best interest of the public. When the extremists are of the *p-type*, it plays *Welfare*; when the extremists are of the *m-type*, it takes security measures. In the separating equilibrium, we do not observe attacks from *p-type* extremists, though the existence of this equilibrium is ensured by the threat that, should the government opt for defense, the public would respond by withdrawing its support with a high probability.

Counter-intuitively, the probability of public support in response to $\{D, Atk\}$ is positively, rather than inversely, related to the probability of *p-types* playing $\{Atk|D\}$: as the latent threat of attacks decreases, so does the latent support for the government, should they attack. That is, *p-type* extremists are able to signal their type to the public by restraining themselves from attacking, which causes the public to withdraw its support from the government. This

relationship reverses in the semi-separating equilibrium, as I'll discuss shortly.

The shift from a pooling to a separating equilibrium occurs when both public support and the probability of an attack decline beyond a threshold. When the extremists are weakened, and the public's threat perception goes down, the government may have a tougher time justifying new security measures. This may explain why governments sometimes offer concessions to extremists when the extremists are substantially weakened. For instance, the 1998 Good Friday Agreement between the UK government and the IRA was reached after the IRA was substantially weakened and infiltrated by British intelligence (Frampton 2016).

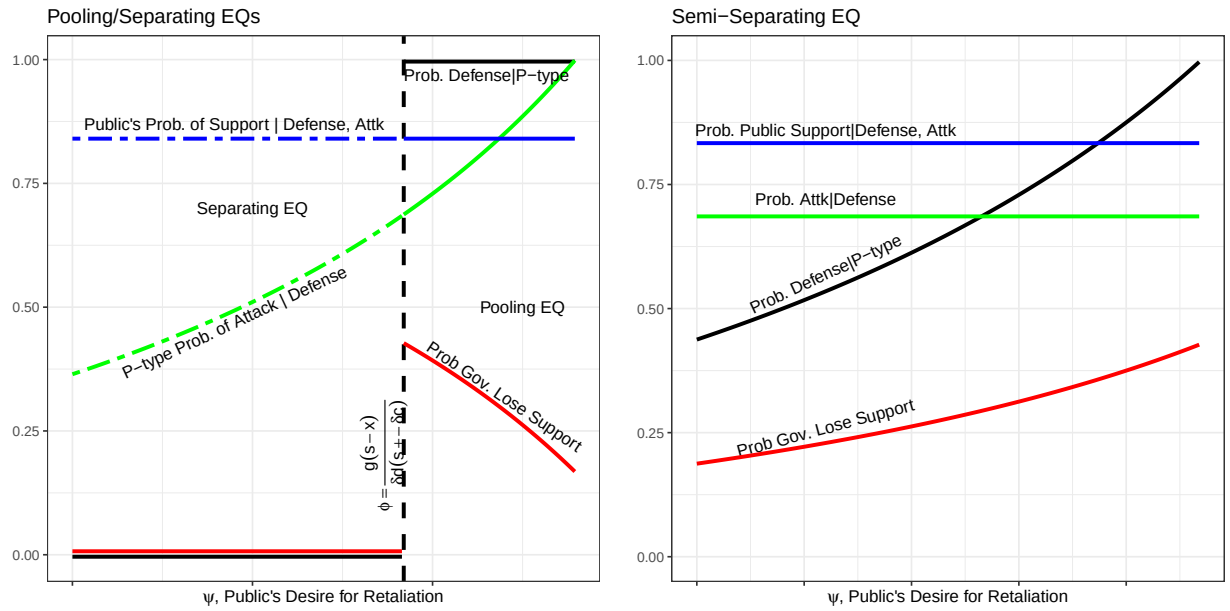
The second subfigure shows the relationship among the same probabilities under semi-separating equilibria, for a set of plausible parameter values. Essentially, the government is enticed to play *Defense* against *p-types* when (a) both the probability of public support, p, and the probability of an attack, γ, are fairly high and (b) one of these probabilities is especially high. The strategies by the public and the *p-type* extremists are inversely related: the extremists increase their probability of attack when the public decreases its support, and vice versa. As public support decreases and the probability of attacks increases, the government acts on its predatory incentive with a lower probability.

The Role of the Democratic Public

Under what conditions can a democratic public constrain a predatory incumbent who may be exaggerating the extremist threat in order to stay in power? To answer this question, the game specifies a sophisticated payoff function for the public, one that accounts for more than the basic needs for safety and well-being.

Value for Retaliation: A key parameter in the public's payoff function reflects its desire for retaliation against the attackers, ψ. Wayne (2023) shows that extremist attacks elicit anger on the part of the public, which in turn evokes a desire for retaliation against the

Figure 3: The Effect of the Good Governance Parameter on Players' Actions and Government Support



Note. Black line shows the probability that the government plays D against *p-types*, green denotes the probability that the *p-type* extremists attack in response to defense, blue is the probability that the public support the government after a game history of $\{D, Atk\}$, and red the probability of a game outcome in which the public does not support the government. Dashed lines indicate unobservable parts of the equilibria, solid lines indicate observable parts. All parameters are set at reasonable fixed values.

perceived attacker (McDermott, Lopez, and Hatemi 2017; Schnakenberg and Wayne 2024). Building on these insights, I incorporate the desire for retaliation into the public's payoff structure and explore its effects. Figure 3 displays the retaliation parameter ψ on the x-axis and the probabilities of interest on the y-axis. Line color and type are as defined in Figure 2.

The left-hand subfigure shows that, holding other parameters constant, as the public develops a stronger desire for retaliation, the parameter space changes enabling a pooling equilibrium, in which the government plays D so as to increase its electoral incentives rather than protect the public from the threat of attacks. Increasing ψ also increases the probability of *p-type* extremists' attack in response to D.

The right-hand subfigure demonstrates the effect of ψ under the semi-separating equilibria. For higher values of ψ, the government can afford to act on its private electoral

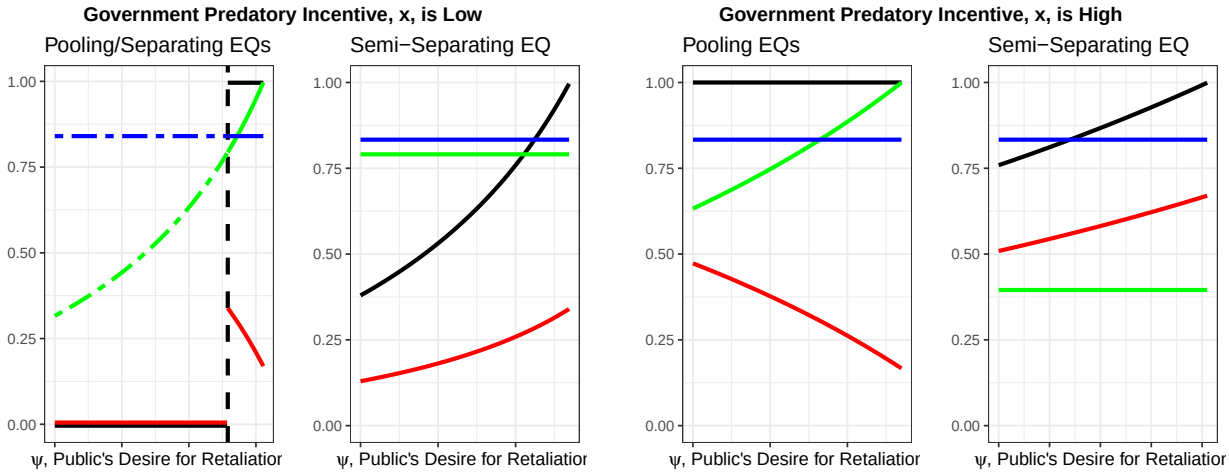
incentive with a higher probability, all the while the public support does not change. Under the semi-separating equilibria, however, the public's calls for retaliation do not affect the probability of an attack by *p-type* extremists. In other words, a predatory government may take advantage of the public's anger and desire for retaliation to advance its electoral goals. A government that prioritizes its own survival may have an interest in whipping the public into a frenzy so as to help justify its private aims. Even-headedness on the part of the public, in contrast, constrains a leader from pursuing predatory incentives.

This analysis explains why, in response to extremist attacks, governments often use emotionally evocative framing. In the case of Turkey (1994–2002), the PKK were labelled "baby killers" and "foreign agents" (Bechev 2022, 28), the latter term favored by other repressive regimes, such as Putin's Russia (Seskuria 2021). Painting the opposition as evil unties the state's hands by providing justification for any policy response. The effectiveness of such framing, however, depends on the availability of a counter-narrative. In 1994–2002 Turkey, the mainstream narrative of the PKK was challenged by the group's own outreach efforts (O'Connor 2021, 111).

A nuance here is that in the semi-separating equilibrium, an increase in ψ also corresponds with an increase in the probability that the government loses the support of the pivotal voter. That is because the probability of the game history D, A, S is the joint probability of these actions and increases with the probability of D: as the government plays D over W with an increased probability, it faces a higher probability that the *p-type* extremists attack and the public plays $\neg S$.

Value for Retaliation, Varying the Government's Predatory Incentive: While all governments want to stay in power, they vary as to the lengths to which they would go in pursuit of this aim. At the extreme, a predatory government might increase its term in office by using (an exaggerated) extremist threat to justify declaring a state of emergency,

Figure 4: The Effect of the Retaliation Parameter, Varying Government Predatory Incentive



Note. Black line shows the probability that the government plays D against *p-types*, green denotes the probability that the *p-type* extremists attack in response to defense, blue is the probability that the public support the government after a game history of $\{D, Atk\}$, and red the probability of a game outcome in which the public does not support the government. Dashed lines indicate unobservable parts of the equilibria, solid lines indicate observable parts. All parameters are set at reasonable fixed values.

restricting opposition and the media, and suspending elections. Most incumbents, however, might stop short of such drastic action.

In the game, an incumbent's value for resorting to predatory measures is modeled via the x parameter. Figure 4 shows the differences in actors' equilibrium strategies in response to a change in the public's value for vengeance, ψ, for low and high values of x. Moving from left to right, the first two subfigures demonstrate the changes in actors' strategies when x is low, that is when the incumbent has low interest in keeping power through the use of predatory measures (e.g., the United Kingdom during the Troubles). In the third and fourth subfigures, x is set to a higher level (e.g., Erdoğan's Turkey, 2003–2012; Israel post October 7, 2023). All other game parameters are fixed at values within their constraints.

The figures show that, for low levels of ψ, governments with low x, either play W against the *p-type* extremists (as part of a separating equilibrium) or mix strategies with a relatively low probability of playing D, γ. Governments with high x, in contrast, either play a pure strategy of D (in a pooling equilibrium) or mix with a high probability of playing D, γ.

For high levels of γ, all types of governments play D against *p-type* extremists with an increasing probability, γ. For incumbents with low x, as ψ increases beyond a certain threshold, ψ^*, the game moves into a pooling equilibrium in which they play a pure strategy of D–the same as incumbents with high x. In the separating equilibria, increases in ψ lead to increases in the probability of D against the *p-type* extremists, γ, albeit the high-x incumbents increase from a higher starting value.

This analysis isolates public outrage as a sufficient condition, under which democratic institutions create incentives for governments—predatory and non-predatory alike—to opt for a military strategy against *p-type* extremists. High values of ψ set up conditions for two scenarios—in both, the public is not any safer from extremist attacks; the latter is also characterized by a risk of democratic backsliding.

In the first scenario, a non-predatory government responds to the public's calls for retribution by implementing inefficient security measures against the *p-type* extremists who would have otherwise accepted concessions. The *p-types* respond by escalating their attacks, which lowers the level of public safety. An example is British Prime Minister Thatcher's interactions with the IRA during the 1981 Hunger Strikes. While Thatcher's intransigent stance[5] during the Hunger Strikes drew much international condemnation, the British public enthusiastically supported their leader. So high was the public outrage with the IRA that even the BBC's tepid efforts to provide comprehensive coverage of events put the Corporation in the crossfire, raising concerns for the safety of its journalists (Savage 2022).[6] Thatcher's tough stance did not, however, stop future IRA attacks, which persisted through the 1990s. Another example of a public-induced inefficient escalation is the case of Turkey

[5]Archival evidence of secret negotiations between PM Thatcher and the IRA during the 1981 Hunger Strikes indicates that Thatcher did not offer any substantive concessions on IRA's demands. The PM's offer was limited to prison visiting, clothing, and parcels (BBC 2011).

[6]The Controller of the regional BBC service in Northern Ireland, James Hawthorne, received credible death threats against his wife and children and, for his protection, was moved to a 'safe house' outside of Belfast for six weeks and given a different car for the purpose of disguise (Savage 2022, 17).

and the PKK between 1994–2002, discussed below.

The second scenario—described in detail in the fourth case study below (Erdoğan vs. Ergenekon 2008–2012)—is a case, in which an angry public does not only induce inefficient escalation, but also sets up conditions for democratic backsliding.

"Good Governance": A key parameter that enters into the public's payoff structure is the value for a "non-predatory" government, g. Empirically, the value of "good governance" may manifest as an anti-corruption sentiment among the public, which may intensify during economic slumps. As discussed in the case study section, the Turkish public viewed the poor economic performance in the mid and late 1990s as the "governance failure and abuse of power," a perception intensified by numerous contemporaneous corruption scandals (Bechev 2022, 23).

Recall, in Figure 2, the x-axis represents g, the "good governance" parameter that proxies the value a player assigns to being ruled by an "honest" government. In the game, an honest government acts in the best interest of the public rather than pursuing its own predatory incentive to stay in power. That is, an honest government plays W when the extremists are of the *p-type* and D when the extremists are of the *m-type*. The parameter g enters the payoff functions of both the public and the extremists, whose value for an honest government are correlated at $\sigma > 0$. The y-axis represents players' actions as previously described.

The equilibria analysis shows that higher values of g correspond to the equilibria space, in which the public withdraws its support from a government that plays D with an increasing probability. This is the case for both the pure strategy (subfigure on the left) and the semi-separating equilibria (subfigure on the right). In the left-hand subfigure, this decrease in public support forces the government to start playing W when the extremists are of the *p-type*. In the semi-separating equilibria, the government also starts playing W with an increasing probability. Thus, close public scrutiny of government decision-making acts

another constraining mechanism on a predatory leader.

A notable nuance here is the strategy of the *p-type* extremists. If the starting point is the left-hand subfigure, where the actors are playing pooling equilibrium, a positive shift in g corresponds to a negative shift in the probability that *p-types* attack in response to the government playing D. This effect is the opposite in the semi-separating equilibrium, where an increase in g results in an increase of the same probability. In other words, the relationship between government concessions and extremist attacks is not uniquely defined, which offers an explanation for an unresolved debate in the literature regarding the relationship between extremist attacks and concessions (cf. Fortna 2015; Thomas 2014). The observed outcome under the separating equilibria is that the government makes policy concessions and the extremists do not attack, which may suggest that concessions reduce attacks. Under the semi-separating equilibria, in contrast, the probability of government concessions is inversely related to the probability of attacks.

Illustrating the Equilibria: A Case Study of Turkey, 1984–2012

The relationship between political extremism, counter-extremist policies, and public support is mired with endogeneity. Governments may respond to extremist threat with counter-insurgency measures or by seeking a political middle ground via concessions. Extremists may escalate their threat in response to either strategy. The public may reward a more militant or a more conciliatory government response, depending on the the broader political temperature (hawkishness, trust in government). The existence of multiple equilibria, each characterized by different—sometimes opposing—dynamics among player strategies, precludes straightforward statistical analysis, such as correlating government support with insurgents attacks.

Nonetheless, the model provides a lens for analyzing historical events and zeroing in on the

causal mechanics that led to specific outcomes, especially in contexts when inter-relationships among actor strategies evade clean causal identification. I illustrate the model's insights by applying it to the case of Turkey between 1984–2012.

This case fits within the model's assumptions. The start of the period marks Turkey's return to democracy after the 1980 military coup as well as the rise of the Kurdistan Worker's Party (PKK) insurgency. The end of the period is characterized by Erdoğan's consolidation of power, the uncovering and crackdown against Ergenekon, and a move towards authoritarianism. During the period under study, Turkey met the institutional criteria specified by the model, in that its government required public support to stay in power.

Both the PKK and Ergenekon are *p-type* extremist groups. Though its stated goals have morphed in response to local and global political context, the PKK has always advocated for political and economic rights and protections for the Kurdish minority, as well as democratic protections for the broader Turkish population (O'Connor 2021; Bechev 2022).[7] Ergenekon consisted of ultranationalist rogue security operatives who opposed Erdoğan's religious reform and aimed to maintain secularism in politics. The group was uncovered in 2007 and linked to a series of high-profile assassinations of politicians and public figures going back to the 1990s.

Based on actor strategies, I further divided the case into four sub-periods that correspond to different game equilibria: PKK in 1984–1993 (democratization, through the death of President Özal), PKK in 1994-2002 (before Erdoğan became Prime Minister), PKK in 2003–2012 (Erdoğan's rule, prior to consolidation of power), and Ergenekon in 2008-2012.

The advantage of a temporal case study is that it allows for a 'loosely controlled' analysis

[7]For example, the PKK sought to overturn the infamous Law 2932 banning the Kurdish language from the political domain, passed by the junta in 1983 (O'Connor 2021, 110). The group released a manifesto clarifying its objective as a struggle for independence and democracy and not to maximize casualties. An emphasis on targeted violence, which contrasted with the indiscriminate brutality of the junta, was a part of the PKK strategy to gain popular support within the broader Kurdish society and attract recruits (O'Connor 2021, 109).

Table 2: Parameter Values for the Case of Turkey

Period	*Gov predatory motive, x*	*Public's concern w/corruption, g*	*Public's vengeance, ψ*	*Outcome*
1984–1993 PKK	Low	Low	Low	Semi-separating EQ: All players mix.
1994–2002 PKK	Low	High	Medium	Pooling EQ: G plays D, other players mix.
2003–2012 PKK	High	Low	Low	Semi-separating EQ: All players mix.
2008–2012 Ergenekon	High	Medium	High	Pooling EQ: G plays D Ergenekon plays Atk; The Public mixes.

of the variation in the outcome in response to changes in a small number of parameters, while holding all else constant. The main sources of variation in the exogenous parameters are the incumbent's predatory motive, x, the public's concern with government's corruption, g, and the public's value of retaliation against the extremists, ψ. All other parameters are assumed fixed throughout the entire time period. Table 2 summarizes the values of the key parameters and the corresponding outcomes for each time period. In the next section, I focus on how equilibria change under different values of g and ψ. I discuss the effect of changing x in the subsequent section.

Turkey–PKK, 1984–1993

During this period, the actor strategies correspond to a semi-separating equilibrium. Both the government and the extremists alternate between a show of force and negotiations, while the insurgents grow their support with the public. After a three-year period of post-coup rule, the junta lifted a ban on political parties and a newly formed Motherland Party (ANAP) won the parliamentary majority (O'Connor 2021, 107). The ANAP's creator, Turgut Özal, became Turkey's Prime Minister. The period also marked the transition of the PKK from a

leftist student movement that pursued its goal via available conventional means to a small rural-based guerrilla army that, in 1984, launched a military insurgency.

In retrospect, the period is viewed as a "missed opportunity with regard to the Kurshish issue" (Bechev 2022, 20). Özal was open to negotiations with the PKK; he himself identified as part Kurdish; his party ran on a big-tent platform drawing support from a diverse constituency, and included several prominent (assimilated) Kurds (O'Connor 2021, 19). Nonethess, these were the early years in the PKK fight: it took time before the insurgents gained support and signaled their strength. Once it did, in the early 1990s, the state shifted towards offering a greater mix of concessions, yet any progress halted after Özal's untimely death.

The Government: In the model's parlance, Özal played a mixed stratgy vis-a-vis the PKK: on one hand, he launched a military campaign and enacted the Emergency Law curtailing civil liberties in Kurdish provinces; on the other hand, he also initiated negotiations with the PKK and provided several important concessions.

Initially, the Turkish government did not view the early attacks by the PKK as a major security concern (O'Connor 2021, 110). Nonetheless, the state responded to the attacks by sending reinforcements to the region and establishing military outposts in sympathetic villages. The resulting casualties and arrests caused some locals to denounce the PKK out of fear of the state. By March 1986, the state escalated by reviving the system of local militia units, the Village Guards, with the goal of fighting the insurgents (O'Connor 2021, 122). To infiltrate the PKK, the state passed a 'Repentant Law' that enticed insurgents to switch sides (O'Connor 2021, 111). The state further escalated by imposing the Emergency Law in July 1987 in the so-called OHAL region—a term used to refer to the areas affected by the insurgency. Per this legislation, a state-appointed governor of the OHAL region had extraordinary powers, such as to censor the press, remove judges, and limit the right

to assembly. The same law later facilitated mass expulsion of the Kurds from the area (O'Connor 2021, 129). The boundaries on the OHAL region expanded as necessary to reflect the insurgency spread (O'Connor 2021, 129).

As the insurgents gained in strength over time, and the state support in the region waned, the military campaign was supplemented with concessions, such as lifting the bans on Kurdish language and the celebration of *Newroz* (Kurdish New Year) in 1993 (Bechev 2022, 20). Prior to his death of a heart attack in 1993, Özal was prepared to negotiate with the militants; part of this initiative was the insurgent-declared ceasefire. This carrots-and-sticks approach corresponds to a mixed government strategy consistent with a semi-separating equilibrium, depicted in the second subfigures of Figures 2 and 4.

The Extremists: During this time period, the PKK plays a mixed strategy. Its attacks are targeted and localized to the rural areas. From their bases in rural mountainous villages near the Iraqi and Syrian border, the PKK launched a series of attacks on the army barracks and gendarmerie bases. Despite major setbacks, the PKK was able to survive the winter of 1984-85, maintaining a foothold in a few villages, which proved its resilience and inspired further support from the local population (O'Connor 2021, 113).

For the remainder of the 1980s, the PKK grew in numbers and territorial control, increasing the intensity and frequency of attacks, and inflicting growing casualties on the Turkish forces. Referred to as *serhidan*, the early 90s was the period of uprising akin to the Palestinian *intifada* (O'Connor 2021, 178). Despite the growing numbers and support, it chose not to escalate to a full-scale conflict. Attacks are intended as a show of strength aimed at gaining support; at the same time, the insurgents demonstrated restraint.

For example, after three successive years of week-long violence during the celebration of the Kurdish New Year between 1990–92, the 1993 celebration were peaceful, after the PKK leader, Öcalan, declared a ceasefire (O'Connor 2021, 111). Another example of restraint

is the clear separation between the rural and urban campaigns. Despite substantial urban support, the PKK limited direct fighting to the rural areas, drawing on urban centers for recruits, resources, and "the terrain for cultivating a legal Kurdish movement" (O'Connor 2021, 180).

The Public: The public also plays a mixed strategy. The initial attacks in 1984 helped rally the PKK Kurdish supporters, as well as the broader non-Kurdish sympathizers. Aware that the mainstream Turkish media was unlikely to report on the attacks in a flattering light, the PKK engaged in a concerted campaign to spread their message to nearby communities, taking advantage of their existing networks of supporters as well as working to expand these networks (O'Connor 2021, 111). The campaign, along with the localized nature of the fighting, were successful at counter-acting the government's narrative, to keep the public's value of the retribution parameter, ψ, low.

The early 1990s saw the rise of civilian support for the PKK (Marcus 2007, 175) and a decrease in the support for the government (O'Connor 2021, 174). The insurgents probed the extent of public support by calling on civilians to attend public funerals for the fallen insurgents. These funerals drew thousands of civilian mourners in the rebel-controlled region (Marcus 2007, 175). Another focusing event for popular demonstrations was the celebration of the Kurdish New Year. In three consecutive years, between 1990–1992, *Newroz* was a week-long event, "in which tens of thousands of Kurds took to the streets throughout the region to taunt the military and shout PKK slogans" (Marcus 2007, 175).

Turkey–PKK, 1994–2002

During this period, the actor strategies correspond to the pooling equilibrium. This case is another example of inefficient escalation driven by the pivotal voter's calls for retribution. In 1994, the Turkish economy experienced the worst recession in decades—GDP contracted

4.7 percent, and the lira depreciated from 15,000 to 38,000 relative to the US dollar (Bechev 2022, 15). By the end of the 1990s, the national budget ran a deficit of 11.7 percent of GDP (Bechev 2022, 15). Between 1994–99, President Süleyman Demirel, oversaw a series of weak parliamentary coalitions involving leftists and Islamists. Meanwhile, the government extended the term for mandatory military service from 15 to 18 months, so it could maintain the 250,000 troops fighting the PKK in the southeast provinces (Bechev 2022, 16). By the end of the decade, casualties reached more than 30,000, including 5,828 Turkish officials, 5,390 civilians, and 19,789 PKK fighters, while around 4,000 villages had been evacuated, with an estimated cost for the conflict of 86 billion USD (Bechev 2022, 27).

The Government: In the model's terminology, during this time period, the government played a pure strategy of *Defense*. Prime Minister Tansu Çiller (June 1993—March 1996) believed in a military solution to the insurgency, abandoning Özal-era peace overtures and stepping up military operations in the southeast as well as northern Iraq (Bechev 2022, 27). The culmination of the military campaign was the arrest of PKK leader Öcalan after Turkey had threatened Syria, who aided in Öcalan's protection, with a ground invasion. The military campaign was coupled with arrests of Kurdish politicians and activists, such as Leyla Zana, the first Kurdish woman to be elected to parliament. The gendermerie's unofficial intelligence arm, JITEM, targeted Kurdish activists and politicians with assassinations (Bechev 2022, 27). The Turkish state also delegated some extra-judicial killings and violence to Hezbollah (O'Connor 2021, 164–171).

The Extremists: In response to the government strategy of *Defense*, the PKK played a pure strategy of *Attack*. Relying on external support from Syria, Lebanon, and Iraq, Kurdish militants launched cross-border attacks, setting fire to government buildings and killing Turkish soldiers (Bechev 2022, 27). While in the 1980s, the PKK violence was restricted to the southeast, in the 1990s, the PKK strategy expanded to western Turkey, where it included

periodic bombing campaigns and clashes arising from the escalation of small demonstrations. Öcalan's arrest did not lead to the dissolution of the PKK: though its guerilla membership declined to just several thousand, its command structure of Cemil Bayık, Duran Kalkan, and Murat Karayılan, stayed intact (Bechev 2022, 28; O'Connor 2021, 184).

The Public: The public's strategy was mixed, though the pivotal voter had shifted against the state. As is often the case during economic downturns, the public's anti-corruption sentiment, g, increased, as it blamed the government for the economic woes. Added fuel to the outrage was a series of poorly timed corruption scandals, ranging form funneling of loans to cronies to connections between high-ranking politicians and organized crime (Bechev 2022, 23).

Beyond the battlefield, the PKK and the state also fought—with mixed success—an informational war over the control of the dominant narrative, so as to influence the public's desire for retribution ψ. On one hand, the escalation and the accompanying increase in casualties, resulted in calls for revenge by some of the public. Public hostility was fueled by the state-led media campaign that labeled the PKK as "baby killers" and agents of the state's foreign enemies and played up the organization's links to drug trafficking. The PKK's external support from Syria, Greece, Iran, and Russia gave traction to the anti-Kurdish sentiment (Bechev 2022, 28). On the other hand, the period corresponds to a large-scale internal displacement from the southeast to the west of Turkey, shifting demographics and voting preferences in key urban areas. The displacement was not only due to the fighting, but also to the Southeast Anatolia Project of building a chain of dams and hydroelectric plants. In all, over 3 million people migrated between the early 1980s to the late 1990s (Bechev 2022, 28).

The flow of Kurdish migrants to big cities fueled broader national awareness of the conflict previously restricted to the southeast (O'Connor 2021, 192). The PKK, despite lacking

territorial control in urban centers, built support among the new migrants though service provision, e.g., after-school Kurdish-language classes, arbitration of disputes (O'Connor 2021, 171). Whereas previously the Kurdish issue was restricted to local identity-based parties, it was increasing its national appeal, through a faction within the SHP, a left-wing party that was in and out of parliamentary governing coalitions.[8]

Turkey–PKK, 2003–2012

During this period, the players are in a semi-separating equilibrium yet again: the government and the extremists alternate between escalation and negotiations, while the public supports limited concessions. This period was shaped by two broad shifts in the political landscape: the rise of the Islamist parties and Turkey's EU ambition. Amid the 1990s economic slump, several high-profile corruption scandals surrounding the center-right parties, pushed voters to seek political alternatives, making the Islamist party, Refah, an indispensable parliamentary coalition partner. In particular, in the 1994 local elections, Refah won across Turkey, and one of its young charismatic leaders, Recep Tayyip Erdoğan, won the mayorship of Istanbul (Bechev 2022). Islamists drew support from religious Kurds by focusing on the religious identity as a uniting factor between Turks and Kurds (Bechev 2022, 30).

Two consecutive economic crises, in 2000 and 2001, cemented popular support in favor of the Islamist party. By 2000, Turkey's external debt reached 60 percent of GDP, driving away foreign investment and causing depreciation of the lira, unemployment, and falling wages. By 2001, bank bailouts drove public debt to 74 percent of GDP. Eager to punish the mainstream parties that had held power since the 1990s, in 2002 voters fled to the AKP, an offshoot of Refah, led by Erdoğan, who ran on the platform of Europeanization and economic reform.

[8]In 1990, several members of the SHP splintered and formed the People's Workers Party (HEP) that advocated for schooling and broadcasting in Kurdish. The HEP was banned in 1993, as were its successors.

In December 2002, EU leaders announced that Turkey could start membership talks by 2004, conditional on meeting the political criteria of functional democracy, rule of law, human rights and rights of minorities. To satisfy these criteria, Turkey initiated a series of laws, including a constitutional amendment to abolish the death penalty. One of the biggest remaining issues was finding a resolution to the Kurdish issue.[9] The political reality was that Erdoğan's AKP had to respond to at least some of the Kurdish demands both in order to satisfy EU conditions and to gain support in the southeast provinces (Bechev 2022, 61).

The Government: The government played a mixed strategy. In contrast to his predecessor, Ergodan also incorporated a conciliatory approach towards the Kurds, stressing shared religious ideology and even admitting previous wrong-doings by the state. He lifted the 1987 Emergency Rule that gave the state special powers in the OHAL region. Per 'local languages' law of 2002, for the first time, the media was allowed to broadcast in Kurdish language. Erdoğan also initiated some exploratory contacts with the PKK, with secret negotiations starting in Oslo in 2008. At the same time, he also took efforts to limit Kurdish movement through the 2006 amendments to the Law on Fighting Terrorism which restricted media publication of pro-Kurdish materials and public display of PKK imagery (Bechev 2022, 65). By the end of 2011, 4,000 people were arrested in relation to this law, including thousands of Kurdish politicians, journalists, academics, and human rights activists (Yesil 2014, 163).

The Extremists: The extremists played a mixed strategy. Öcalan's arrest in 1999 brought a de-escalation of military conflict and a simultaneous rise of the nonviolent part of the Kurdish movement. Weakened by heavy casualties and the loss of their leader, the PKK announced another unilateral ceasefire, with the remaining guerillas withdrawing from Turkey to Iraq (Candar 2020, 119). In 2004, the PKK ended the ceasefire with the aim of bringing

[9]The other two primary issues were the Cypress dispute and civil-military relations.

the state to the negotiating table. The group moderated their demands, replacing the call for independence with that for democratic autonomy (Bechev 2022, 64). Another ceasefire was announced in April of 2009 after the what looked like progress in negotiations, with the fighting resuming by December. A lasting settlement akin to the Good Friday Agreement appeared within reach as late as 2013 (Bechev 2022, 129–130), which gave Erdoğan plenty of time to consolidate power.

The Public: The pivotal voter played a mixed strategy. The nationalist voters opposed excessive concessions to the PKK, while the Turkish mainstream opposed any devolution of power. Given the delicate balance between offering the Kurds some degree of recognition and concessions, especially on language, Erdoğan's AKP fared well against Kurdish nationalist candidates in the east and southeast in the 2004 election by using an incrementalist approach (Bechev 2022, 62).

Turkey–Ergenekon, 2008–2012

This case corresponds to a pooling equilibrium: the government takes a tough stance against the alleged extremists, and the public initially supports the government. The difference between this case and the earlier examples of pooling equilibria (Thatcher and the IRA or Turkey in 1994–2002) is that Erdoğan has a high value of x and uses the situation to consolidate his hold on power.

As the negotiations with the PKK were ongoing, a different extremist organization, Ergenekon, came to the fore of Turkish politics in 2007. Ergenekon was a clandestine network of rogue security operatives and ultranationalists with a secularist agenda whose alleged goal was to assassinate Erdoğan along with several other prominent political figures in order to "prod the military to step in and ultimately drive the AKP out of power" (Bechev 2022, 122; see also Demiroz and Kapucu 2012, Yesil 2014).

Ergenokon's opposition to Erdoğan's was rooted in a long tradition of secularism espoused by Kemal Ataturk, a figure akin to the Founding Father in Turkey who was much revered by the military and many mainstream Turks (Bechev 2022, 16). To Ergenekon, Erdoğan's religious ideology was an affront to religious reform, in particular the hotly contested law allowing women to wear headscarves in public institutions (Bechev 2022, 16). Ergenekon's goal, in other words, was grounded in the policy of maintaining secularism to politics. In regards to the Kurdish issue, the group opposed recognition or political concessions to the Kurds and supported a military solution.

The Government: Erdoğan's response was swift and decisive: over 500 individuals with supposed connections to this organization were arrested between 2008–2009. Later, the prosecutors linked Ergenekon to several high-profile murders and attacks going back to the mid-1990s (Bechev 2022, 122). Seemingly cognizant of the dual-purpose of security measures, posited by the theoretical model, Erdoğan "took full advantage" of Ergenekon as a way to arrest and prosecute "the pillars of the old regime" and his political opposition, including high-ranking military officers, security service operatives, bureaucrats and organized crime bosses (Bechev 2022, 122). The leader of the CHP, the left-wing political opposition to Erdoğan, decried these arrests accusing Erdoğan of "a civilian coup" (quated in Bechev 2022, 122). In 2011, under the guise of the Ergenekon investigation, police raided the offices of a secularist news organization, OdaTV, and charged its investigative journalists with extremism. The ensuing cases were condemned by the West, including organizations such as Reporters without Borders (Bechev 2022, 122).

The Extremists: Ergenekon's strategy was to play *Atk*: investigations linked the group to numerous assassinations of public figures going back to the 90s, with some recent prominent cases in 2006-2007 (Bechev 2022, 122).

The Public: The public plays a mixed strategy. It is uninformed regarding the type of the extremists. Initially the public is convinced by government's narrative (the retributions parameter, ψ, is high) and calls for justice against the terrorists. However, the public starts holding the government's claims to higher scrutiny—the "good governance" parameter, g, increases—once the anti-terrorism campaign expands to engulf less likely suspects such as the media and political opposition.

The cracks in public's trust of Erdoğan is evident in that the CHP gained electoral margins in the 2009 local elections and the 2011 parliamentary elections. Despite losing seats in the Grand National Assembly, Erdoğan's AKP won an unprecedented third term and came just short of winning 50 percent of the national vote, its best result yet (Bechev 2022, 125).

Though scholars disagree on the exact timing of Turkey's democratic backsliding, Erdoğan's consolidation of power in his second and future terms made subsequent time periods no longer suitable for analysis using the model. As the EU negotiations came to an impasse over the issue of Cyprus, so did the peace talks with the PKK. Erdoğan had used the previous decade to curtail the power of the military whose insistence on separation between religion and the state were at odds with his party's platform. Erdoğan did not stop at military reform: through a series of additional reforms, he was able to subvert the courts, the independent media, and any remaining checks on his power.

Democratic Backsliding

In several of the equilibria, the government faces a probability of losing this support (shown in red in Figures 2–4). Though the game is agnostic regarding what precisely happens in such outcomes, a loss of the support of the pivotal voter in a democratic regime implies two possible scenarios: government turnover or democratic backsliding. In the game, the govern-

ment accrues a benefit of x whenever it plays D. This parameter is justified as an additional level of security in office that results from the dual-purpose of defense: for example, a larger internal police force may both help detain the extremists and anti-government protesters.

In the case of losing the support of the public, the government will have to decide if the accrued level of x is sufficient to consolidate its power without the pivotal voter. The last case study of Erdoğan's interactions with Ergenekon illustrates the causal mechanisms of how a government may consolidate its power under the guise of security provision.

Erdoğan's predecessors were constrained from pursuing x. In Özal's case, the constraining factor was threat from the military in the aftermath of the junta rule, while the post-Özal Prime Ministers presided over weak coalitionary governments. In contrast, Erdoğan started his term with a strong parliamentary majority.[10] Armed with a popular mandate, Erdoğan had the opportunity to subvert Turkey's democratic institutions. In retrospect, he also had the will to do so. Hence, the government's value of x during Erdoğan's rule is coded as high.

Notably, Erdoğan did not have the favorable conditions to use the threat of the PKK to consolidate his power between 2003–2008. Erdoğan came to power as a populist anti-establishment candidate who vowed to fix the economy and break up the old political elites, widely perceived as corrupt and ineffective. That is, at the start of his rule, he enjoyed high levels of public trust (the anti-corruption sentiment, g, was low). At the same time, the public's desire for retribution against the PKK, ψ, was low, a result of the PKK outreach campaign as well as the recognition that guaranteeing minority rights were a precursor for the widely popular Europeanization.

Both of these parameter values took on different values with respect to Ergenekon. At the new of the Ergenekon investigation, the value of public desire for retribution ψ, was

[10] A key institutional feature that allowed for this was a 10 percent threshold required to get seats in the parliament. The result was that Erdoğan's party got an overwhelming 69 percent of the seats having won only 34.42 percent of the vote, while 46 percent of the population voted for parties that did not clear the threshold (Bechev 2022, 53).

high: the public called for justice against the rogue security officials, which gave Erdoğan the pretext for arrests of security officials. Once the list of alleged perpetrators expanded to include journalists and public officials, and drew domestic and international condemnation, the public outrage was replaced with suspicion. At that point, the damage was already done, with many of Erdoğan's enemies removed from their positions, weakened, and disgraced.

In summary, the Turkish leader got to live the best of both worlds—mix in some W to the the PKK, which yielded gains at the polls, while decimating his political opposition and the military establishment by implementing D against Ergenekon.

Conclusion

Democratic governments are designed to serve the public, but this design does not require the public to be fair or reasonable. The model isolates the conditions, under which the informational asymmetry in this relationship may reduce, rather than advance the public's core goals of safety, or even enable the demise of democratic institutions. The key mechanisms at work are the public's value of retribution, the public trust in government, and the government's self-interested motive of power consolidation. The first of these mechanisms—the public's value of retribution—is a sufficient condition for inefficient escalation and decreased public safety. The presence of all three set up conditions for democratic backsliding.

In other words, an emotional public motivated by retaliation may voluntarily give up its electoral constraint on predatory executives by supporting repressive policies and restrictions on civil rights and privacy. In fact, when the public clamors for blood, even a non-predatory government may opt for a military strategy so as to safeguard its electoral prospects.

I illustrated the game's insights using the case of Turkey (1983–2012) and a shorter vignette of Thatcher's interactions with the IRA during the 1981 Hunger Strikes. Additional examples include Russia's democratic backsliding in the early to mid 2000s, as it waged

it's Second War in Chechnya, Israel's interactions with the PLO, as well as the currently unfolding events in Gaza.

The results highlight the importance of media independence and public access to alternative media sources. In contrast to autocracies, democratic regimes do not guarantee that the government have an informational advantage over the alternative voices, although the threat of extremism may lead to forms of censorship even democratic regimes.[11] To paraphrase an old adage, while battles may rage on the front lines, the real victories are often secured in the pages of the most widely read tabloids, where public opinion is shaped and sentiments sway.

[11] E.g., Britain's infamous 'broadcasting ban' of 1987 (Savage 2022).

References

Abrahms, Max. 2006. "Why Terrorism Does Not Work." *International Security* 31(2): 42–78.

Allon, Gideon. 2024. "'No doubt' Netanyahu preventing hostage deal, charges ex-spokesman of Families Forum." *The Times of Israel* April 26. Accessed: October 16, 2024. https://www.timesofisrael.com/no-doubt-netanyahu-preventing-hostage-deal-charges-ex-spokesman-of-families-forum/.

Arce, Daniel G., and Todd Sandler. 2007. "Terrorist Signalling and the Value of Intelligence." *British Journal of Political Science* 37(4): 573–586.

Arce, Daniel G., and Todd Sandler. 2010. "Terrorist Spectaculars: Backlash Attacks and the Focus of Intelligence." *Journal of Conflict Resolution* 54(2): 354–373.

BBC. 2011. "Margaret Thatcher 'negotiated with IRA'." December 30. Accessed: October 16, 2024. https://www.bbc.com/news/uk-northern-ireland-16366413.

Bechev, Dimitar. 2022. *Turkey Under Erdoğan*. Yale University Press.

Berrebi, Claude, and Esteban F Klor. 2006. "On Terrorism and Electoral Outcomes Theory and Evidence from the Israeli-Palestinian Conflict." *Journal of Conflict Resolution* 50(6): 899–925.

Brooks, Clem, and Jeff Manza. 2013. "A Broken Public? Americans' Tesponses to the Great Recession." *American Sociological Review* 78(5): 727–748.

Bueno De Mesquita, Ethan. 2005. "Conciliation, Counterterrorism, and Patterns of Terrorist Violence." *International Organization* 59(01): 145–176.

Bueno de Mesquita, Ethan, and Eric S Dickson. 2007. "The Propaganda of the Deed:

Terrorism, Counterterrorism, and Mobilization." *American Journal of Political Science* 51(2): 364–381.

Candar, Cengiz. 2020. *Turkey's Mission Impossible: War and Peace with the Kurds.* Lexington Books.

CBC. 2014. "Viktor Yanukovych Calls for Truce with Ukraine Protest Leaders." February 19. Accessed: October 16, 2024. `https://www.cbc.ca/news/world/viktor-yanukovych-calls-for-truce-with-ukraine-protest-leaders-1.2542830`.

Conrad, Courtenay R., Sarah E. Croco, Brad T. Gomez, and Will H. Moore. 2018. "Threat Perception and American Support for Torture." *Political Behavior* 40: 989–1009.

Cunningham, Kathleen Gallagher. 2011. "Divide and Conquer or Divide and Concede: How Do States Respond to Internally Divided Separatists?" *American Political Science Review* 105(2): 275–297.

Demiroz, Fatih, and Naim Kapucu. 2012. "Anatomy of a Dark Network: the Case of the Turkish Ergenekon Terrorist Organization." *Trends in Organized Crime* 15: 271–295.

Denemark, David. 2012. "Trust, Efficacy and Opposition to Anti-Terrorism Police Power: Australia in Comparative Perspective." *Australian Journal of Political Science* 47(1): 91–113.

Di Lonardo, Livio. 2017. "The Partisan Politics of Counterterrorism: Reputations, Policy Transparency, and Electoral Outcomes." *Political Science Research and Methods* , 1–17.

Dragu, Tiberiu. 2011. "Is There a Trade-off Between Security and Liberty? Executive Bias, Privacy Protections, and Terrorism Prevention." *American Political Science Review* 105(1): 64–78.

Dragu, Tiberiu. 2017. "The Moral Hazard of Terrorism Prevention." *Journal of Politics* 79(1): 223–236.

Dragu, Tiberiu, and Mattias Polborn. 2014. "The Rule of Law in the Fight against Terrorism." *American Journal of Political Science* 58(2): 511–525.

Findley, Michael G., and Joseph K. Young. 2012. "Terrorism and Civil War: A Spatial and Temporal Approach to a Conceptual Problem." *Perspectives on Politics* 10(2): 285–305.

Fisk, Kerstin, Jennifer L. Merolla, and Jennifer M. Ramos. 2019. "Emotions, Terrorist Threat, and Drones: Anger Drives Support for Drone Strikes." *Journal of Conflict Resolution* 63(4): 976–1000.

Fortna, Virginia Page. 2015. "Do Terrorists Win? Rebels' Use of Terrorism and Civil War Outcomes." *International Organization* 69(3): 519–556.

Frampton, Martyn. 2016. *The Long March: The Political Strategy of Sinn Féin, 1981-2007*. Springer.

Huddy, Leonie, and Stanley Feldman. 2011. "Americans Respond Politically to 9/11: Understanding the Impact of the Terrorist Attacks and Their Aftermath." *American Psychologist* 66(6): 455.

Huddy, Leonie, Stanley Feldman, Charles Taber, and Gallya Lahav. 2005. "Threat, Anxiety, and Support of Antiterrorism Policies." *American Journal of Political Science* 49(3): 593–608.

Kydd, Andrew H., and Barbara F. Walter. 2006. "The Strategies of Terrorism." *International Security* 31(1): 49–80.

Lapan, Harvey E., and Todd Sandler. 1993. "Terrorism and Signalling." *European Journal of Political Economy* 9(3): 383–397.

Lyall, Jason. 2009. "Does Indiscriminate Violence Incite Insurgent Attacks? Evidence from Chechnya." *Journal of Conflict Resolution* 53(3): 331–362.

Malhotra, Neil, and Elizabeth Popp. 2012. "Bridging partisan divisions over antiterrorism policies: The role of threat perceptions." *Political Research Quarterly* 65(1): 34–47.

Marcus, Aliza. 2007. *Blood and belief: The PKK and the Kurdish fight for independence.* NYU Press.

McDermott, Rose, Anthony C. Lopez, and Peter K. Hatemi. 2017. "'Blunt Not the Heart, Enrage It': The Psychology of Revenge and Deterrence." *Texas National Security Review* 1(1): 68–88.

Merolla, Jennifer L., and Elizabeth J. Zechmeister. 2009. "Terrorist Threat, Leadership, and the Vote: Evidence from Three Experiments." *Political Behavior* 31: 575–601.

New York Times. 2024. "Biden says there is 'every reason' to believe Netanyahu is prolonging the war for political reasons." June 4. Accessed: October 16, 2024. `https://www.nytimes.com/live/2024/06/04/world/israel-gaza-war-hamas`.

O'Connor, Francis. 2021. *Understanding Insurgency: Popular Support for the PKK in Turkey.* Cambridge University Press.

Overgaard, Per Baltzer. 1994. "The Scale of Terrorist Attacks as a Signal of Resources." *Journal of Conflict Resolution* 38(3): 452–478.

Pape, Robert A. 2003. "The Strategic Logic of Suicide Terrorism." *American Political Science Review* 97(3): 343–361.

Ritter, Emily Hencken. 2014. "Policy Disputes, Political Survival, and the Onset and Severity of State Repression." *Journal of Conflict Resolution* 58(1): 143–168.

Ritter, Emily Hencken, and Courtenay R. Conrad. 2016. "Preventing and Responding to Dissent: The Observational Challenges of Explaining Strategic Repression." *American Political Science Review* 110(1): 85–99.

Savage, Robert J. 2022. *Northern Ireland, the BBC, and Censorship in Thatcher's Britain.* Oxford University Press.

Schnakenberg, Keith E., and Carly N. Wayne. 2024. "Anger and Political Conflict Dynamics." *American Political Science Review* , 1–16.

Schneider, Friedrich, Tilman Brück, and Daniel Meierrieks. 2015. "The Economics of Counterterrorism: A Survey." *Journal of Economic Surveys* 29(1): 131–157.

Seskuria, Natia. 2021. "Why Putin Is Obsessed with 'Foreign Agents'." *Foreign Policy* . Seotember 13, `https://foreignpolicy.com/2021/09/13/vladimir-putin-foreign-agents-law-russia/`.

The Atlantic. 2024. "Netanyahu is prolonging Gaza war to stay in power, Palestinian FM claims." March 14. Accessed: October 16, 2024. https://www.theatlantic.com/international/archive/2024/09/netanyahus-other-war/679786/.

Thomas, Jakana. 2014. "Rewarding Bad Behavior: How Governments Respond to Terrorism in Civil War." *American Journal of Political Science* 58(4): 804–818.

Walter, Barbara F. 2006*a*. "Building Reputation: Why Governments Fight Some Separatists but not Others." *American Journal of Political Science* 50(2): 313–330.

Walter, Barbara F. 2006*b*. "Information, uncertainty, and the decision to secede." *International Organization* 60(1): 105–135.

Wang, Yuhua, and Bruce J. Dickson. 2022. "How Corruption Investigations Undermine Regime Support: Evidence from China." *Political Science Research and Methods* 10(1): 33–48.

Wayne, Carly N. 2023. "Terrified or Enraged? Emotional Microfoundations of Public Counterterror Attitudes." *International Organization* 77(4): 824–847.

Yesil, Bilge. 2014. "Press Censorship in Turkey: Networks of State Power, Commercial Pressures, and Self-Censorship." *Communication, Culture and Critique* 7(2): 154–173.

Appendix: Game Solution

The game is solved using the Perfect Bayesian equilibrium (PBE) solution concept.

Separating Equilibrium 1

The game has a separating equilibrium, in which the *p-type* extremists never attack, the *m-type* extremists always attack, the government plays W when the extremists are of the *p-type* and D when the extremists are of the *m-type*, and the public supports the government in all observable outcomes.

Notice that for the *m-type* extremists, the strategy of $\neg Atk$ is strictly dominated. Denote *m-type*'s expected utility from attacking and not attacking as $U_{ME}(Atk)$ and $U_{ME}(\neg Atk)$, accordingly. Since $d > \delta d > r$, *m-type*'s expected utilities from attacking a government that plays D are $U_{ME}(Atk|D, S) = \delta d$ if the public supports the government and $U_{ME}(Atk|D, \neg S) = r + \delta d$ if the public does not support the government. In contrast, *m-types* expected utilities from not attacking are $U_{ME}(\neg Atk|D, S) = 0$ and $U_{ME}(\neg Atk|D, \neg S) = r$. As $\delta d + r > \delta d > r > 0$, the *m-type* extremists always play Atk after observing D, irrespective of the strategy of the public.

Likewise, *m-type*'s expected utilities from attacking a government that plays W are $U_{ME}(Atk|W, S) = d$ if the public supports the government and $U_{ME}(Atk|W, \neg S) = r + d$ if the public does not support the government. In contrast, *m-types* expected utilities from not attacking are $U_{ME}(\neg Atk|W, S) = 0$ and $U_{ME}(\neg Atk|W, \neg S) = r$. As $d + r > d > r > 0$, the *m-type* extremists will always play Atk after observing W. Hence, *m-type* extremists always play Atk, irrespective of the strategies of the government or the public.

Next, let us consider the strategy of the *p-type* extremists. Notice that the *p-type* extremists prefer not to attack as long as the government plays W. Denote *p-type*'s expected utility from attacking and not attacking as $U_{PE}(Atk)$ and $U_{PE}(\neg Atk)$, accordingly. Assuming the

government plays W, the four possible expected utilities to consider are:

1. *P-type* extremists' expected utility from Atk should the public support the government, $U_{PE}(Atk|W, S) = g - d + b$;

2. *P-type* extremists' expected utility from Atk should the public not support the government, $U_{PE}(Atk|W, \neg S) = -d + b$;

3. *P-type* extremists' expected utility from $\neg Atk$ should the public support the government, $U_{PE}(\neg Atk|W, S) = g + b$;

4. *P-type* extremists' expected utility from $\neg Atk$ should the public not support the government, $U_{PE}(\neg Atk|W, \neg S) = b$;

The restriction $d > g$ leads to the ordering of the above payoffs, such that $g + b > b > g - d + b > b - d$, with both payoffs from $\neg Atk$ outweighing either of the payoffs from Atk. That is, *p-type* extremists' best response to W does not depend on whether the public supports the government: as long as the government plays W, *p-type* extremists never attack.

Assuming the government plays D, the four possible expected utilities to consider are:

1. *P-type* extremists' expected utility from Atk should the public support the government, $U_{PE}(Atk|D, S) = \delta d$;

2. *P-type* extremists' expected utility from Atk should the public not support the government, $U_{PE}(Atk|D, \neg S) = g + \delta d$;

3. *P-type* extremists' expected utility from $\neg Atk$ should the public support the government,
   $U_{PE}(\neg Atk|D, S) = 0$;

4. *P-type* extremists' expected utility from $\neg Atk$ should the public not support the government, $U_{PE}(\neg Atk|D, \neg S) = g$;

Comparing the four expected utilities, we note that the *p-type* extremists' best response depends on the public's choice. Specifically, the *p-type* extremists' expected utility from Atk outweighs that from $\neg Atk$ if the public does not support the government upon observing D. In contrast, the *p-type* extremists' expected utility from $\neg Atk$ outweighs that from Atk if the public supports the government upon observing D.

Suppose θ_1 is the public's posterior belief that the extremists are of the *p-type* upon observing a game history of D, A. In a separating equilibrium, the public is able to infer the extremists' type from the history of the game. Since *p-type* extremists never attack, this posterior belief, θ_1, equals to 0. In this case, the public's expected utility from supporting the government, $U_P(S|D, Atk) = \rho b - \delta k + \sigma g + \psi$ outweighs its expected utility from not supporting the government $U_P(\neg S|D, Atk) = \rho b - \delta k$.

Therefore, the public will play S, which means that the *p-type* extremists' best response to D is $\neg A$. Hence, upon observing $D, \neg A$, the public would infer that the extremists must be of the *p-type*. Denote this posterior belief as $\theta_2 = 1$.

Fully specifying this equilibrium requires defining the public's beliefs for each of the other game histories, on and off the equilibrium paths: $\{W, Atk\}$, $\{W, \neg Atk\}$. Denote these believes, respectively, as θ_3 and θ_4. Then, after observing a history of $\{W, \neg Atk\}$, the public can update its belief as $\theta_4 = 1$. In this case, the public's expected utility from supporting the government is $U_P(S|W, \neg Atk) = \rho b + \sigma g$, which is greater than their expected utility from not supporting the government, $U_P(\neg S|W, \neg Atk) = \rho b$. In other words, the public will always support the government if the government chooses W and the extremist group plays $\neg Atk$.

Conversely, if the government plays W and the extremist group attacks, the public would infer that the extremists must be of the *m-type*, i.e. $\theta_3 = 0$. In this scenario, the public's expected utility from supporting the government is $U_P(S|W, Atk) = -k - \sigma g$, which is less than their expected utility from not supporting the government, $U_P(\neg S|W, Atk) = -k$.

Therefore, the public will not support the government if the government plays W and the extremist group attacks.

Now let's consider the decision of the government. If the extremists are of the *p-type*, then the government's expected utility from D is $U_G(D|\neg A, \neg S) = x$ and its expected utility from playing W is $U_G(W|\neg A, S) = s$. Hence, the government will play W when the extremists are of the *p-type* as long as $s \geq x$, i.e. as long as it values the support of the pivotal voter over its power to suppress the opposition. If the extremists are of the *m-type*, the government's expected utility from playing D is $U_G(D|A, S) = s + x - \delta c$, and $U_G(W|A, \neg S) = 0$. This means that the government will play D against *m-type* groups as long as $s + x \geq \delta c$.

This produces the first game equilibrium, $EQ1$:

$$\begin{cases} S_G = \{W, D\}; \\ S_P = \{S, \neg S, \neg S, S\}, \theta_1 = 0, \theta_2 = 1, \theta_3 = 0, \theta_4 = 1; \\ S_{PE} = \{\neg Atk, \neg Atk\}; \\ S_{ME} = \{Atk, Atk\}; \\ s > x, s + x \geq \delta c;\ d > g > \delta d;\ d > r;\ \sigma g > \delta k - \psi \end{cases}$$

The payoffs:

$$\begin{cases} U_G(EQ1) = s + (1 - \alpha)(x - \delta c) \\ U_P(EQ1) = \rho b + \sigma g + (1 - \alpha)(\psi - \delta k) \\ U_{PT}(EQ1) = g + b \\ U_{MT}(EQ1) = \delta d \end{cases}$$

This is a separating equilibrium—the public is able to infer the extremist's type from the observable actions of the government and the extremists. The government prioritizes public

safety and welfare over its own private incentives and reaches an acceptable policy solution with the *p-type* group. The only time the government enacts counterterrorism measures is in response to an imminent threat from militant extremist groups. The public supports the government in all observed outcomes.

Pooling Equilibrium

The game also has a pooling equilibrium, in which the government always plays D, both types of extremists attack, and the public plays a mixed strategy of supporting the government with probability p.

Consider the extremists' decision following a game history, in which the government plays D. As established in the above analysis, the *p-type* extremists' best response at this decision node depends on the public's strategy: they will play Atk as long as the public plays S with such a probability p that ensures that *p-type*'s expected utility from Atk outweighs their expected utility from $\neg Atk$, or $U_{PE}(Atk|D,p) > U_{PE}(\neg Atk|D,p)$. Otherwise, the *p-type* extremists will play $\neg Atk$. Specifically, *p-type* extremists' expected utility from attacking equals to $U_{PE}(Atk|D,p) = p\delta d + (1-p)(g+\delta d) = g + \delta d - pg$, whereas the best they can get from $\neg A$ equals $U_{PE}(\neg Atk|D,\neg S) = g$. Hence, the *p-type* extremists will play Atk as long as the public responds to a game history of D, A by playing S with the probability $p < \frac{\delta d}{g}$.

As before, suppose θ_1 is the public's posterior belief that the extremists are of the *p-type* upon observing D, A. Then the public's expected utility from playing S at this decision node is $U_P(S|D,A) = \theta_1(-\delta k - \sigma g + \psi) + (1-\theta_1)(\rho b - \delta k + \sigma g + \psi)$, and its expected utility from playing $\neg S$ equals $U_P(S|D,A) = -\theta_1 \delta k + (1-\theta_1)(\rho b - \delta k)$. Setting the two expected utilities to equal and solving for θ_1, we obtain that the public will play S at this decision node as long as $\theta_1 < \frac{1}{2} + \frac{\psi}{2\sigma g}$, and $\neg S$ if $\theta_1 > \frac{1}{2} + \frac{\psi}{2\sigma g}$. The condition of $\theta_1 = \frac{1}{2} + \frac{\psi}{2\sigma g}$ is the indifference condition at which the public can play S with any probability p.

From the above analysis, we know that the public's decision depends on θ_1. There are

two cases, in which the *p-type* extremists will play Atk following D that satisfy the condition $p < \frac{\delta d}{g}$. The first case is that of $\theta_1 > \frac{1}{2} + \frac{\psi}{2\sigma g}$, in which the public always plays $\neg S$ at this decision node. The second case is that of $\theta_1 = \frac{1}{2} + \frac{\psi}{2\sigma g}$, which allows the public to play S at this decision node with probability $0 < p < \frac{\delta d}{g}$.

Next, we have to check the stability of the government's strategy. As previously demonstrated, should the government play W, both types of extremists have strictly dominated strategies that do not depend on the choice of the public: the *p-type* extremists always prefer to play $\neg Atk$, whereas the *m-type* extremists always prefer to play Atk. Hence, if the government plays W, the two types of extremists play different strategies, which allows the public to infer their type by updating its believes.

As before, assume that θ_3 is the public's posterior belief that the extremists are of the *p-type* upon observing a game history of $W, \neg Atk$, and θ_4 is the public's belief that the extremists are of the *p-type* upon observing a game history of W, Atk. Because the *p-type* extremists never attack at this decision node, the public is able to update its belief to $\theta_3 = 0$ and $\theta_4 = 1$. Based on these posterior believes, the public will never support the government upon observing W, Atk as its expected utility from $\neg S$, $U_P(\neg S|W, Atk) = -k$ is greater than its expected utility from S, which is $U_G(\neg S|W, Atk) = -k - \sigma g$. Conversely, the public will always support the government upon observing a game history of $W, \neg Atk$, as its expected utility from S, $U_P(S|W, \neg Atk) = \rho b + \sigma g$ outweighs that from $\neg S$, which is $U_G(\neg S|W, \neg Atk) = \rho b$.

Case P1. $S_P(D|A) = \neg S, \theta_1 > \frac{1}{2} + \frac{\psi}{2\sigma g}$: Given the above-specified strategies of the two extremist types and the public, the government's deviating from D to W, in response to *p-type* extremists, would result in a payoff of $U_G(W|\neg Atk, S) = s$, which outweighs its payoff from $U_G(D|Atk, \neg S)$ as long as $s \geq x$. Hence, the government can improve its payoff from 0 to s by unilaterally deviating from D to W. Therefore, the government's strategy is not

stable.

Case P2. $S_P(D|A) = p, p < \frac{\delta d}{g}, \theta_1 = \frac{1}{2} + \frac{\psi}{2\sigma g}$: In this case, the government's expected utility from D is $U_G(D|Atk, p) = p(s - \delta c) + x$, whereas its expected utility from deviating to W is $U_G(W|\neg Atk, S) = s$. By comparing the two expected utilities, we obtain that the government will play D as long as the public plays S with a probability $p > \frac{s-x}{s-\delta c}$. To ensure that p is less than or equal to 1, let $\delta c < s$ and $x > \delta c$. Hence, this condition yields a stable strategy for the government.

Next, we must specify the public's choice at the decision node following a game history $D, \neg A$. As this node is never reached as part of this equilibrium, the public's decision at this node depends on its belief, θ_2 that the extremists are of the *p-type*. Because the strategies of the other players do not depend on the public's choice at this decision node, the public can technically play S at this decision node with any probability. For example, if $\theta_2 = 1$, then the public would always play $\neg S$. If however, $\theta_2 = 0$, then the public would always play S. Hence, the public's choice at this decision node is $S_P(D, \neg A) = \cdot$ and its belief is $\theta_2 \in [0, 1]$.

Finally, the public's posterior belief $\theta_1 = \frac{1}{2} + \frac{\psi}{2\sigma g}$ that the extremists are of the *p-type* upon observing a game history of D, A, must be consistent with its prior belief α. By Bayesian updating, $\theta_1 = \alpha$ or $\alpha = \frac{1}{2} + \frac{\psi}{2\sigma g}$.

This produces the second equilibrium class, EQ2:

$$\begin{cases} S_G = \{D, D\}; \\ S_P = \{p, \cdot, \neg S, S\}, \\ \theta_1 = \alpha = \frac{1}{2} + \frac{\psi}{2\sigma g}, \theta_2 \in [0,1], \theta_3 = 0, \theta_4 = 1; \\ S_{PE} = \{Atk, \neg Atk\}; \\ S_{ME} = \{Atk, Atk\}; \\ \frac{s-x}{s-\delta c} < p < \frac{\delta d}{g}; \\ s > x,\ s \geq \delta c;\ \delta c \leq x;\ d > g > \delta d;\ d > \delta r;\ \sigma g > \psi \end{cases}$$

The payoffs:

$$\begin{cases} U_G(EQ2) = p(s - \delta c) + x \\ U_P(EQ2) = p(\alpha\sigma g + \sigma g + \psi + b) - \alpha(\sigma g + \psi) - \alpha(\sigma g + \rho b) + \rho b - \delta k \\ U_{PT}(EQ2) = g + \delta d - pg \\ U_{MT}(EQ2) = r + \delta d - pr \end{cases}$$

This is a pooling equilibrium, in that the strategies of the informed actors (the government and the extremists) reveal no additional information regarding the type of the extremists to the public. In this equilibrium, the government prioritizes defense measures over policies aimed to enhance the welfare of the public, both extremist groups attack, and the public supports the government with a positive probability p. This equilibrium is possible under two conditions: (1) a high probability that the extremists are of the policy type ($\alpha = \frac{1}{2} + \frac{\psi}{2\sigma g}$) (2) the government values the private benefit of counter-terrorist measures more than public safety, $\delta c \leq x$. In this equilibrium, both types of extremist groups execute attacks against the government and the public, even though the government could have averted the attacks by *p-types*. Ironically, in this equilibrium, the government's emphasis on defense results in a less

safe outcome for the public. Another counter-intuitive insight is that, although the public is aware that the government may act in a predatory manner (both x and δc are public knowledge), it still supports the government with a positive probability p. The public's willingness to support the government, despite a high probability that the government is acting in a predatory manner, is driven by the value of the public's hawkishness ψ relative to the public's appreciation for a non-predatory government σg. As the ratio $\frac{\psi}{\sigma g}$ approaches 1, the public's preference for "being tough on extremists" takes over even as they are aware that the extremists are likely of the policy type and that their government is likely acting on its own predatory incentive. The existence of this equilibrium suggests that predatory governments may be able to take advantage of the public's hawkishness to enact self-serving policies, such as limiting civil liberties and restricting privacy in the name of security.

Semi-Separating Equilibrium 1

The game has a semi-separating equilibrium, in which the government plays D regardless of the extremist group's type, the *p-type* extremists play a mixed strategy—they attack with probability ϕ and do not attack with probability $1-\phi$, and the *m-type* extremists always attack. In this equilibrium, the public supports the government upon observing D, Atk with probability p and does not support the government with a probability $1-p$. Upon observing a game history $D, \neg Atk$, the public does not support the government.

P-type's Indifference Condition: Consider the game histories, in which the government plays D. The above analysis established that the *p-type* extremists are indifferent (can play Atk with any probability) between their two choices if the public responds to Atk by playing S with probability $p = \frac{\delta d}{g}$ and responds to $\neg Atk$ by playing $\neg S$.

First, check the stability of the government's strategy. Let ϕ denote the probability that the *p-type* extremists attack in response to D. The government's expected utility from

playing D against the *p-type* extremists is $U_G(D|\phi, \{p, \neg S\}) = \phi p(s+x-\delta c)+x$, whereas it's expected utility from playing W is $U_G(W|\neg Atk, S) = s$. Hence, the government's indifference condition is:

$$\phi = \frac{g(s-x)}{\delta d(s-\delta c)}.$$

The government will play D at this decision node, as long as $\phi \geq \frac{g(s-x)}{\delta d(s-\delta c)}$ and W otherwise.

Case G1. $\phi > \frac{g(s-x)}{\delta d(s-\delta c)}$**:** The next step is to derive the public's posterior beliefs following each game history. We already know that, upon observing a game history D, A, the public will play a mixed strategy as along as its posterior belief that the extremists are of the *p-type* is $\theta_1 = \frac{1}{2} + \frac{\psi}{2\sigma g}$.

By Bayes' Law, consistency between the public's prior and posterior beliefs requires:

$$\frac{\psi + \sigma g}{2\sigma g} = \frac{\alpha\phi}{\alpha\phi + 1 - \alpha}.$$

Solving for ϕ:

$$\phi = \frac{(1-\alpha)(\sigma g + \psi)}{\alpha(\sigma g - \psi)}.$$

Next, consider the public's belief at the decision node following a game history $D, \neg A$. Since *m-type* extremists always attack, upon observing no attack following D the public will update its belief and conclude that the extremists must be of the *p-type*. Hence, $\theta_2 = 1$.

The final step is to define the public's beliefs at the decision nodes that are not reached as part of this equilibrium. As before, suppose θ_3 is the public's posterior belief that the extremists are of the *p-type* upon observing a game history of $W, \neg Atk$, and θ_4 is the public's belief that the extremists are of the *p-type* upon observing a game history of W, Atk. Because the *p-type* extremists never attack at this decision node, I assume that $\theta_3 = 0$ and $\theta_4 = 1$.

Based on these posterior believes, the public will never support the government upon observing W, Atk as its expected utility from $\neg S$, $U_P(\neg S|W, Atk) = -k$ is greater than its expected utility from S, which is $U_G(\neg S|W, Atk) = -k - \sigma g$. Conversely, the public will always support the government upon observing a game history of $W, \neg Atk$, as its expected utility from S, $U_P(S|W, \neg Atk) = \rho b + \sigma g$ outweighs that from $\neg S$, which is $U_G(\neg S|W, \neg Atk) = \rho b$.

The result is a semi-separating equilibrium, in which the government plays D against both types of the extremists. The *p-type* extremists attack with probability $\phi = \frac{(1-\alpha)(\sigma g+\psi)}{\alpha(\sigma g-\psi)}$. The *m-type* extremists always attack. Following D, Atk, the public plays S with probability $p = \frac{\delta d}{g}$, and $\neg S$ with probability $1 - p$. And at the decision node $D, \neg Atk$, the public does not support the government.

This produces the third equilibrium, EQ3:

$$\begin{cases} S_G = \{D, D\}; \\ S_P = \{\frac{\delta d}{g}, \neg S, \neg S, S\}, \\ \theta_1 = \frac{1}{2} + \frac{\psi}{2\sigma g}, \theta_2 = 1, \theta_3 = 0, \theta_4 = 1; \\ S_{PE} = \{\phi = \frac{(1-\alpha)(\sigma g+\psi)}{\alpha(\sigma g-\psi)}, \neg Atk\}; \\ S_{ME} = \{Atk, Atk\}; \\ \phi > \frac{g(s-x)}{\delta d(s-\delta c)}; \\ s > x,\ s \geq \delta c;\ \delta c \leq x;\ d > g > \delta d;\ d > \delta r;\ \sigma g > \psi \end{cases}$$

The payoffs:

$$\begin{cases} U_G(EQ3) = \left(\frac{\delta d}{g}(s - \delta c) + x\right)(\alpha\frac{(1-\alpha)(\sigma g+\psi)}{\alpha(\sigma g-\psi)} + 1 - \alpha) + \alpha x(1 - \frac{(1-\alpha)(\sigma g+\psi)}{\alpha(\sigma g-\psi)}) \\ U_P(EQ3) = \alpha\frac{(1-\alpha)(\sigma g+\psi)}{\alpha(\sigma g-\psi)}\frac{\delta d}{g}(-\delta k - \sigma g + \psi) - \alpha\frac{(1-\alpha)(\sigma g+\psi)}{\alpha(\sigma g-\psi)}(1 - \frac{\delta d}{g})(\delta k) + (1-\alpha)(\rho b - \delta k + \sigma g + \psi) \\ U_{PE}(EQ3) = g \\ U_{ME}(EQ3) = r + \delta d(1 - \frac{r}{g}) \end{cases}$$

Semi-Separating Equilibrium 2

The game has another semi-separating equilibrium, in which the government always defends against *m-type* extremists, but plays a mixed strategy when dealing with *p-type* extremists: it plays D with probability γ and W with probability $1-\gamma$. The *p-type* extremists never attack in response to W, but play a mixed strategy in response to D: attack with probability ϕ. The public never supports the government after $D, \neg Atk$ or $W, \neg Atk$, but plays a mixed strategy in response to D, Atk: supports the government with probability p, and removes the government with probability $1-p$.

Case G2. $\phi = \frac{g(s-x)}{\delta d(s-\delta c)}$: Recall that the government's indifference condition is $\phi = \frac{g(s-x)}{\delta d(s-\delta c)}$. Also recall that, upon observing a game history D, A, the public will play a mixed strategy as along as its posterior belief that the extremists are of the *p-type* is $\theta_1 = \frac{1}{2} + \frac{\psi}{2\sigma g}$.

By Bayes' Law,

$$\frac{\psi + \sigma g}{2\sigma g} = \frac{\alpha\gamma\phi}{\alpha\gamma\phi + 1 - \alpha}.$$

Solving for $\gamma\phi$:

$$\gamma\phi = \frac{(1-\alpha)(\sigma g + \psi)}{\alpha(\sigma g - \psi)}.$$

The public's posterior beliefs at other decision nodes are defined as previously: $\theta_2 = 1$, $\theta_3 = 0$, $\theta_4 = 1$.

As a result, we can summarize the second semi-separating equilibrium, $EQ4$:

$$\begin{cases} S_G = \{\gamma = \frac{(1-\alpha)(\sigma g+\psi)}{\alpha\phi(\sigma g-\psi)}, D\}; \\ S_P = \{\frac{\delta d}{g}, \neg S, \neg S, S\}, \\ \theta_1 = \frac{1}{2} + \frac{\psi}{2\sigma g}, \theta_2 = 1, \theta_3 = 0, \theta_4 = 1; \\ S_{PE} = \{\phi = \frac{(1-\alpha)(\sigma g+\psi)}{\alpha\gamma(\sigma g-\psi)}, \neg Atk\}; \\ S_{ME} = \{Atk, Atk\}; \\ \phi = \frac{g(s-x)}{\delta d(s-\delta c)}; \\ s > x,\ s \geq \delta c;\ \delta c \leq x;\ d > g > \delta d;\ d > \delta r;\ \sigma g > \psi \end{cases}$$

The payoffs:

$$\begin{cases} U_G(EQ4) = \alpha\gamma(\phi p(s-\delta c)+x) + \alpha s(1-\gamma) + (1-\alpha)(p(s\delta c)+x) \\ U_P(EQ4) = \alpha\gamma\phi(p\psi - p\sigma g - \delta k) + (1-\gamma)(\rho b + \sigma g) + (1-\alpha)(p\sigma g + p\psi + \rho b - \delta k) \\ U_{PE}(EQ4) = g + b - \gamma b \\ U_{ME}(EQ4) = r + \delta d(1 - \frac{r}{g}) \end{cases}$$

This is a semi-separating equilibrium. This equilibrium is enabled by several conditions: (1) the public's high posterior belief that the extremists are of the policy type ($\frac{1}{2} + \frac{\psi}{2\sigma g}$), (2) the government positive utility from implementing counterterrorism measures, even after the cost of possible terrorist attacks ($\psi > \sigma g$), and (3) the public's hawkishness (high value of ψ relative to σg). Under this equilibrium, the government always implements counterterrorism measures when the extremist group is of the militant type. When the extremist group is of the policy type, the government implements counterterrorism with probability γ and implements policy change the rest of the time (with probability $1-\gamma$).

Separating Equilibrium 2

Finally, the game has another separating equilibrium, in which the *p-type* extremists never attack, the *m-type* extremists always attack, the government plays W when the extremists are of the *p-type* and D when the extremists are of the *m-type*, and the public supports the government in all observable outcomes. This separating equilibrium is different from EQ1 in the off-equilibrium-path strategies and beliefs. As I show below, these strategies arise when the *p-type* extremists do not attack in response to D with a high enough probability ϕ.

Case G3. $\phi = \frac{s}{\frac{\delta d}{g}(s+x-\delta c)}$**:** In this case the government strictly prefers to play W in response to *p-type* extremists. This equilibrium is supported as long as the public's beliefs are: $\theta_1 = \frac{1}{2} + \frac{\psi}{2\sigma g}$,$\theta_2 = 1$,$\theta_3 = 0$, and $\theta_4 = 1$.

More formally, the second separating equilibrium, $EQ5$ can be summarized as:

$$\begin{cases} S_G = \{W, D\}; \\ S_P = \{\frac{\delta d}{g}, \neg S, \neg S, S\}, \theta_1 = \frac{1}{2} + \frac{\psi}{2\sigma g}, \theta_2 = 1, \theta_3 = 0, \theta_4 = 1; \\ S_{PE} = \{\phi = \frac{(1-\alpha)(\sigma g+\psi)}{\alpha(\sigma g-\psi)}, \neg Atk\}; \\ S_{ME} = \{Atk, Atk\}; \\ \phi < \frac{s}{\frac{\delta d}{g}(s+x-\delta c)}; \\ s + x \geq \delta c;\ d > g > \delta d;\ d > \delta r;\ \sigma g > \psi \end{cases}$$

The payoffs:

$$\begin{cases} U_G(EQ5) = s + (1-\alpha)(x - \delta c) \\ U_P(EQ5) = \rho b + \sigma g + (1-\alpha)(\psi - \delta k) \\ U_{PT}(EQ5) = g + b \\ U_{MT}(EQ5) = \delta d \end{cases}$$