

Transitional Justice Against Agents of Repression and the Threat of Regime Change

Genevieve Bates, Monika Nalepa, Ben Konstan and Jordi Vasquez

The University of Chicago

This version August 30, 2020

Abstract

We investigate the phenomenon of authoritarian backlash following the prosecution of agents of the former authoritarian regime after its democratic transition. Our model separates criminal transitional justice targeting low-ranking agents of repression from the prosecutions of leaders, who issue orders to repress. We show that targeting low-ranking agents of repression increases their incentives for supporting coup leaders planning authoritarian take-overs. We test our model's predictions using the Global Transitional Justice Dataset, which reports prosecutions over time and includes information on the rank of defendants. This information allows us to construct an original independent variable: the proportion of leaders that stood trial as a percentage of all criminal trials against former perpetrators, what we call the "skewness of justice." We demonstrate that although coup attempts are not affected by the skewness of justice, successful coups are. More detailed case studies of Argentina and Egypt illustrate our argument and the mechanism at play.

Prepared for Presentation at the (Virtual) American Political Science Association in September 2020.

The authors are grateful to Ipek Cinar, Viivi Jarvi, Ken Krmoyan, Moksha Sharma, Karina Cheung, Zikai Li, and Ojashwe Sakopta. for stellar research assistance and for funding their work to the National Science Foundation, the Center for International Social Science Research, and the Pearson Institute for Global Conflict at the University of Chicago. All mistakes are the authors' responsibility.

Bear in mind that they would not have been able to do what they did without others to assist them. Nor would they have attempted to come here, except in the hope of being rescued by those same people, who have come not to help them, but in belief that if you acquit those who are responsible for the greatest evils, they themselves will have virtual impunity for what they did and might wish to do so in the future

Lysias in *Against Erasthenes* (Todd et al., 2000a)

1 The Retributive Justice Dilemma

How does criminal transitional justice affect democratic stability? For centuries, new democracies have taken steps to hold members of ancien régimes and their agents of repression accountable. In ancient Athens, following the rule of the Thirty Tyrants (a pro-Spartan oligarchy installed in Athens after its defeat in the Peloponnesian War), several measures were taken against the Tyrants themselves and their collaborators. *Dokimasia* was a screening procedure aimed at vetting who among the Athenian citizens had collaborated with the Thirty. Citizens proven to have collaborated could not hold public office (Todd et al., 2000b). A similar fate awaited those who had served in the cavalry of the “Thirty.” To verify if a citizen had been a calvarist, Athenians consulted the *sainidion*, a register of all calvarists (Todd et al., 2000b). Finally, there were criminal prosecutions. The Thirty themselves and their supporting council of Eleven were prosecuted and, in most cases, sentenced to death, but of their 3000 supporters, all were amnestied and allowed to seek refuge outside of Athens, with one exception: “If they had killed another man with their own hands” (Todd et al., 2000a).

The choice made by the Athenians to prosecute more severely those fulfilling orders than those issuing them is highly unusual in the broader context of transitional justice. More often than not, the rank-and-file are actually spared responsibility for carrying out even the most egregious human rights violations on behalf of the authoritarian regime. Any disproportionality between rank and accountability is usually due to the imperfect implementation of transitional justice.

For instance, following the US War of Independence, loyalist Tories were prosecuted for their collaboration with the British Army. However while collaborators who were captured and hanged tended to be closer to the lower echelons of Patriot military hierarchy, the resource-rich, like Benedict Arnold, who was responsible for considerably more damage done to the revolutionaries, were able to flee to Europe (Elkins and McKittrick, 1995; Brumwell, 2018). Following the Civil War, in 1866, Congress proposed a measure for systematically dealing with spies and collaborators. The law required that persons who wished to carry out “certain occupations” (including the practice of law) subscribe to an oath that they had never “offered aid to the rebellion.” The measure was however, promptly overturned with the Supreme Court striking down the provisions dealing specifically with the disbaring of former members of the Confederate government from practicing as attorneys.¹

Even so, a couple of low-ranking spies who were captured by sheer coincidence were hanged almost immediately.² This last example, however, serves rather as an illustration of “victor’s justice” than of transitional justice.

For a more contemporary example, consider the article by Capoccia and Pop-Eleches (2020), who examine the implementation of denazification policies (mostly purges and lustration) in the three ally-occupied zones in Germany. Because allies’ views on the appropriate severity of transitional justice varied from one zone to another, rank-and file agents of repression in one zone (the American zone) were punished more severely than leaders issuing orders in other zones (French and, to some extent, British).

In many instances, however, the reason it is so hard to prosecute rank and file agents of repression is the principle of non-retroactivity (*nullum crimen sine lege* or “no crime without a law”), a rule of law standard that constrains new democracies and makes it hard

¹*Ex parte Garland*, 71 U.S. 333 (1866).

²The first was Owen Dodd of Little Rock, who along with identification demanded by Union soldiers produced a notebook with encrypted information about Union troop strength in Little Rock. He was arrested, convicted by a court-martial of being a spy, and sentenced to execution by hanging on January 8, 1864. The second case was that of John Yates Beall, apprehended at Niagara Falls for spying and participation in a Lake Erie raid. Even though his attorney defended him claiming that Beall was a Confederate naval officer acting under orders, he was convicted and despite numerous petitions to pardon him, hanged on February 24, 1865.

to prosecute anyone for conduct that was not only legal under the previous constitutional framework, but indeed, encouraged. In light of the above discussion, it is—at least from a legalistic point of view—easier to prosecute order-givers.

The second reason why agents of repression may be spared accountability before transitional justice tribunals is simply because they are innocent. Thus far, we have been taking for granted the willingness of these agents to fulfill orders to perform repression. Yet this need not be the case. According to (Masri, 2017, page 47), Ben Ali “relied heavily on his police forces to quell the populace” because he could not depend on the army that “did not have the power nor political will to intervene on behalf of the regime.” Masri goes on to explain how on numerous occasions, Rachid Ammar-Alis chief of staff refused orders to fire on protesters during the Arab Spring. Ultimately, the army developed a reputation for being an ally and supporter of the demonstrators. All the violent crackdowns on protesters were carried out by security forces. The security forces were also better rewarded and respected than the military by the dictator.³

Perhaps it was the reluctance of the army to follow brutal orders that spared them the harsh transitional justice that focused on Ali’s secret police.

The final reason for avoiding criminal prosecutions of agents of repression is practical: where following the demise of the ancien régime, the military apparatus remains strong well into the democratic era, engaging in transitional justice can antagonize state security forces. For example, after the democratic transition in Argentina, the military was fragmented, but maintained strict control of the state coercive apparatus. Shortly after the October 1983 elections, democratic President Raul Alfonsín signed Decree No. 158, which “ordered the prosecution of all the members of the first three military juntas for the human rights atrocities committed since the 1976 coup” (Human Rights Watch, 2001). Despite growing discontent from the military, several court decisions and policy initiatives from the new government left

³Masri (2017) claimed that “resentment had built up among the forces after Ben Ali directed leadership purges in the 1990s, replacing both senior and junior military officers whom he accused of having Islamist leanings and allowing the police to attempt to exert authority over the armed forces”

mid and lower ranking military officers open to being tried. Outrage over the trials acted as a unifying force, providing incentives for the fragmented military to coalesce behind an anti-transitional justice agenda (Nino, 1998). The discontent came to a head in the spring of 1987, when “middle-level officers led by Col. Aldo Rico revolted in Córdoba and Buenos Aires demanding a full-scale amnesty law” (Human Rights Watch, 2001). In order to stave off a coup, the democratic leadership rapidly ceased the pursuit of criminal trials, and an amnesty remained in place until 2005 (Nino, 1998; Human Rights Watch, 2005).

The Argentinean experience brings into stark relief how hard it is for new democracies to toe the line between democratic stability and accountability. It also underscores the importance of distinguishing between prosecutions focusing on members of the ancien régime who were issuing orders (the leaders) and prosecutions focusing on those following orders (the rank and file) when evaluating the severity of transitional justice prosecutions.

Following a repressive regime and transition to a polyarchy, it is common to address calls for punishing the perpetrators of human rights violations committed in the name of the ancien régime. Such retroactive justice can extend to leaders who issued repressive orders as well as to agents of repression who were tasked with executing those orders. In a world where transitional justice can take either of those forms, the calculus of an agent of repression is not trivial: should he obey an order to repress and risk punishment if the regime should fall? Or should he ignore the order and thus expedite the fall the regime?

In this article, we ask what is the effect of the transitional justice regime on this decision calculus of agents’ of repression. The ultimate answer to this question has implications for democratic stability. We model the Repressive Agent’s Dilemma initially as a game of imperfect information and then as an infinitely repeated game, which we solve for Perfect Markov Equilibria (Gehlbach, 2013). The solutions to these models lead us to expect that harsher transitional justice against the rank and file will enable former authoritarian leaders in their attempts to overthrow democracy by allowing them to more easily mobilize these former rank-and-file officers against the new democratic regime. The normative implication

of our model is thus that, their responsibility for maintaining the ancien régime notwithstanding, even the most culpable of human rights violations rank-and-file members of the authoritarian regime may need to be amnestied in order to avoid authoritarian backlash.

Our paper is organized as follows. The next section presents the two formal models described above—first a baseline model with complete information and next, a Markov-Transition game—and derives their empirical implications. We test our model’s predictions with an originally collected for this purpose dataset of criminal trials of authoritarian perpetrators in the aftermath of democratic transitions. The dataset, which is part of the Global Transitional Justice Dataset, not only codes criminal trials of perpetrators over time, but in addition to the volume of trials, distinguishes between the prosecutions of order-givers and order-takers. Because successful coups are, at the end of the day, relatively rare events, we supplement our empirical investigation with a case study from Egypt. The Egyptian experience vividly illustrates the decision calculus of agents of repression.

2 Modeling the prosecution of agents of repression

In the introduction to the article, we make the argument that on the one hand, transitional justice targeting the rank-and-file decreases their willingness to repress authoritarian resisters and thus can bring about the end of authoritarian rule, contributing to democratization. On the other hand, democracy carries with it constraints that make punishing former agents of repression difficult. These constraints are normative—retroactive justice violates key rule of law principles—and practical—transitional justice decreases the opportunity costs of rank-and-file members to join their former leaders’ attempts at thwarting democracy. In order to determine what the joint effect of these conflicting incentives implies for democratic stability, we need a formal model. Such models have been offered by scholars of democratic transitions and international relations. No one to date, however, has explicitly asked how dealing with

agents of repression will affect democratic stability.⁴

There are two recently published models that explore the relationship between authoritarian leaders and their agents under autocracy.

Dragu and Lupu (2018) use a formal model with incomplete information to show that repression is most likely to hinge on expectations and is the result of a coordination game. Paradoxically, when authoritarian leaders need it most—when dissent against them is at its highest—mobilizing agents of repression is most difficult. While in Dragu and Lupu (2018), the autocrat’s dilemma was limited to the authoritarian period, Tyson (2016) extends the consequences of the autocrat’s actions and the consequences of actions of his agents of repression into the post-authoritarian period, allowing for transitional justice. He models the interaction between a leader and his repressive apparatus in circumstances where the stability of the authoritarian regime is uncertain. The autocrat in these circumstances must compensate his agents of repression to offset their potential of being punished should the regime collapse. Tyson’s model uses the prospect of transitional justice to model repressive agents’ incentives but his theory does not distinguish between agents in different ranks. Our model does this and at the same time is able to make more precise predictions about not only when the former leaders want to stage a coup but also when they can do so successfully.

3 Baseline Model

The baseline model builds on Tyson (2016). There are two players in our baseline model: an order-giver who is the leader of the regime (L) and an order-taker who is a representative member of the regime’s repressive apparatus (R). The order-giver is responsible for:

1. deciding whether to give an order to repress citizens;
2. setting a compensation level to reward an order-taker for successfully carrying out the order

⁴But see non-formal paper by Bates et al. (2020) on how purges of former uniformed members of the ancien régime contribute to crime and especially violent crime.

The order-taker only has the choice between accepting the order and repressing its population and refusing to follow the order.

3.1 Timeline

- In the first stage of the game, L chooses between giving the order to repress ($g = 1$) and not giving the order ($g = 0$).
- If there is no order to repress, the game ends with the regime removed from power.⁵
- If L plays $g = 1$, then in the second stage, he sets the compensation level, $c \in [0, C]$ offered to the order-taker in exchange for executing the order.⁶
- In the third stage, and in response to this information, R decides whether to accept compensation and carry out the order or to refuse compensation and not carry out the order. If R refuses the order, the order-giver punishes him by imposing a cost $\sigma < 0$, the regime falls, and the game ends. There is no transitional justice.
- If the order-taker executes the order of repression, the game enters a fourth stage, where with probability $\gamma \in (0, 1)$ the regime stays in power. Under this scenario, a transfer of compensation between the order-giver and order-taker takes place and the game ends. With probability $1 - \gamma$, the regime is toppled in the fourth stage and the process of transitional justice commences.

To ensure the model's tractability, we assume that the order-giver himself is always punished for his role in the repressive regime as a result of transitional justice.⁷ The order-taker, on the other hand, is held accountable with probability $\varepsilon \in (0, 1)$. ε , thus, represents

⁵Without repression, there is no pretext for transitional justice. This assumption is consistent with others in the literature, including Tyson (2016).

⁶Choosing a compensation level is in line with Tyson (2016).

⁷However, this assumption does not change any of the results or comparative statics, either in the basic model or the extension.

the extent to which transitional justice involves punishing agents of repression alongside their leaders.⁸

3.2 Strategies

A strategy of an order giver is a pair (g, c) where the first element corresponds to the binary choice of ordering ($g = 1$) or not ordering ($g = 0$) repression, whereas the second element corresponds to the level of compensation. Hence the strategy set of the order giver is:

$$S_L = \{1, 0\} \times [0, \infty)$$

A strategy of the order taker is an acceptance region containing all compensation schemes following which R executes the order. Hence the strategy set can be written as:

$$S_R = \{A^C \subset [0, \infty) : c \in A^C \rightarrow R \text{ executes order}\}$$

3.3 Payoffs

To define players' utility functions, we introduce additional notation. Let $P > 0$ be the value of staying in power for the order-giver. Recall that $c \in [0, \infty)$ is the compensation for the order-taker and $\sigma < 0$ is the punishment for failing to execute an order. Finally let $\phi_L < 0$ and $\phi_A < 0$, respectively, define the payoffs from transitional justice when L and A are held accountable.

$$u_L(g, c; A^c) = \begin{cases} g * (\gamma(P - c) - (1 - \gamma) * \phi_L) & \text{if } c \in A^c \\ 0 & \text{if } c \notin A^c \end{cases}$$

$$u_R(g, c; A^c) = \begin{cases} g * (c - (1 - \gamma) * \varepsilon * \phi_R) & \text{if } c \in A^c \\ -\sigma & \text{if } c \notin A^c \end{cases}$$

⁸A possible criticism of this set up, is the assumption that the leader should pay the rank-and-file compensation in the last period of this one shot game, even though it comes to him at a cost and he would be better off forgoing payment. We direct such skeptics to consult Myerson (2008) who structures this trust-dilemma of leaders paying the military apparatus as a repeated game that ensures that the military is still paid in Nash Equilibrium. Since we defer bringing in the repeated game apparatus to the next section, we avoid it here for tractability's sake and assume that even though the authoritarian elite cannot credibly commit to compensation for the military it avoids a coup (Myerson, 2008).

If L plays $g = 0$, both players receive payoff 0 as a baseline.

3.4 Equilibrium Analysis

This is a game with complete although imperfect information. It can be solved for Subgame Perfect Equilibrium through a generalized procedure of backward induction (Kaminski and Nalepa, 2014; Kaminski, 2019). Note first, that in equilibrium only two values of c may be offered. This is because if L knows that R will refuse to take the order, she has no motivation to offer any $c > 0$. Define $c^* = \frac{\sigma - (1-\gamma)\varepsilon\psi_R}{\gamma}$.

This first, non-repressive equilibrium can be constructed as:

$(0, g; [c^*, C])$, where $g = 0, 1$ (Note that since the offer is rejected, it does not matter if the order is actually given or not).

To find the second equilibrium, note that R will only accept a scheme c that compensates him for the punishment he would face from refusing the order. Hence for R to agree to repress we need:

$$EU_R(g, c, A^c) > -\sigma \tag{1}$$

Or alternatively, $c > \frac{\sigma - (1-\gamma)\varepsilon\psi_R}{\gamma}$.

Second, for this compensation scheme to be proposed in equilibrium, the order giver needs to prefer the expected outcome associated with this compensation scheme to the alternative, which is the fall of the regime without transitional justice, i.e.,

$$EU_L(g, c; A^C) > 0 \tag{2}$$

After substituting in the utility function of the order giver, we arrive at: $(P-c)\gamma + (1-\gamma)\phi_L \geq 0$

Moreover, since L 's utility is decreasing in c , in equilibrium, he will choose the lowest c

that satisfies (1). This gives us our first proposition:

Proposition 1 *Let c^* be defined as above. When $P > \frac{\sigma - (1-\gamma)(\varepsilon\phi_A + \phi_L)}{\gamma}$, the profile $(c^*, 1; [c^*, \infty))$ is the unique subgame perfect equilibrium. The probability of a repressive equilibrium is decreasing in ε , the probability of the military apparatus facing transitional justice prosecutions.*

This inequality demonstrates that the decision calculus of the leader results in repression when P , the value gained from staying in power, exceeds the payment he must transfer to his agents of repression, as well as the risk that the leader could face transitional justice. It is then weighed by the probability of successful repression, as the leader only attains P with probability γ .

3.5 Discussion of comparative statics

The clearest way of discussing comparative statics is by defining the discrete probability of a repressive equilibrium: $f(\varepsilon) = P\gamma - \sigma + (1 - \gamma)(\varepsilon\phi_A + \phi_L)$. Repression is in equilibrium when $f > 0$, whereas a peaceful transition occurs when $f < 0$. Increasing f , then, raises the risk of repression.

The first comparative static is intuitive. The more attractive it is for the leader to remain in authoritarian office, the more likely the repressive equilibrium, as f is increasing in P . Similarly intuitive is the comparative static on γ , the probability that the regime survives. An increase in γ increases the probability of the repressive equilibrium.⁹

The last sentence of Proposition 1 contains a somewhat more controversial comparative static: the repressive equilibrium becomes less frequent as transitional justice becomes more likely. To see it is true, notice that f is decreasing in ε . This result follows from taking the derivative with respect to ε in light of the fact that $1 - \gamma > 0$ and $\phi_A < 0$,

The conventional wisdom, however, is that the threat of punishment with transitional justice will urge repressive leaders to hold on to power at all costs, even if those costs

⁹Note, that because the punishment parameters, ϕ_L and ϕ_A are both negative, the negative signs in the definition of f cancel one another out.

entail repression of their citizens (Krcmaric, 2018; Elster, 2004; Nalepa and Powell, 2016). Similarly controversial is the final comparative static, on the magnitude of punishment (here punishment for both the leader, represented by ϕ_L and punishment for the rank-and-file, represented by ϕ_A), according to which the greater the punishment, the lower the probability of a repressive equilibrium.

This is intuitive, because increasing the punishment or its frequency, increases c^* , the cost necessary to get the agent to engage in of repression. When this cost exceeds the benefit of staying in office, violating human rights is simply not be worth it.

In light of this argument, what gives? Is the model we propose an inadequate portrayal of the rank-and-file dilemma or has the existing literature been obscuring how distinct his decision calculus is from that of the leadership? To some extent, there is truth on both sides of this argument. On the one hand, the literature cited above fails to distinguish between order-givers and order-takers in the context of transitional justice. On the the other hand, several assumptions made in our baseline model straitjacket real world relationships as exogenous. For instance, we might expect that the probability of regime survival, γ is dependent on whether the rank and file are choosing to execute orders. We might also wonder why former autocrats and their agents of repression are modeled as passive recipients of transitional justice. In real-world settings, former autocrats may defend themselves from transitional justice, first at the peace negotiations, attempting to extract amnesty and later if the former fails, by staging a coup d'état.

It is this latter possibility that we explore in our extension of the baseline model in the next section.

4 Extension: Transitional Justice and the threat of coup d'état

We now present an extension to this basic game. As explained above, it is unlikely that former agents of repression will just sit back and watch as they and their leaders face accountability and transitional justice. More likely, as the Argentinean case from the introduction illustrates, they will threaten to take up arms against the new democratic regime.

To reflect this possibility, the main change relative to the baseline version is that we now allow the military apparatus to stage a coup following transitional justice. We assume that once out of power, the former authoritarian elite will need the support of the military to successfully launch a coup. Also, the introduction of transitional justice is itself a strategic choice made by a post-authoritarian democratic elite who determine the severity of transitional justice facing the military apparatus.

This new democratic elite faces the trade-off between punishing former repressive agents (to deter them from following inhuman orders) and lowering their opportunity costs so much that they may as well back their former leaders and attempt to re-establish autocratic rule.

Players, States and equilibrium concept

A discrete-time Markov transition game is an infinite horizon game with the following properties:

- There is a set of players, N ;
- There is a set of States, K , where each period, t , has a state $k \in K$ associated with it;
- There is an action set for each player defined as a function of the state, so $A_i(k)$, for $i \in N$
- There is a transition function $g(k^{t+1}|k^t, a^t)$ that specifies for each state k the probability that the game transitions to that state given a previous state and the action profile of

players in that previous state in the previous period. Note, the transition between the states may also be deterministic (This latter property allows us to generate the possibility of exiting and re-entering the authoritarian regime).

- For each player, preferences are represented by the discounted sum of payoffs:

$$\sum_{t=0}^{\infty} \delta^t u_i(k^t, a^t) \tag{3}$$

Our equilibrium concept is the Markov perfect equilibrium, a refinement of the subgame perfect equilibrium. A strategy profile in Markov transition game specifies an action for each player for each state of the game. For a strategy profile to be in Markov perfect equilibrium it has to be the case that a) no player has a unilateral incentive to deviate in any state (a property shared with the subgame perfect equilibrium) and b) strategies are only contingent on the current state. This last property means that we abstract from players' expectations formed on the basis of what happened previously in the game.

Below, we explain the specifics of our Markov transition game where, following transitional justice, democracy can relapse into authoritarianism via coup.

We have three players: an authoritarian leader, L , a representative member of the repressive apparatus, R , and the post-authoritarian democratic elite, E . We also define four states:

1. An authoritarian regime under threat of revolt. This corresponds to the beginning of the baseline model.
2. A democratic regime without transitional justice following a peaceful transition of power, which can be interpreted as corresponding to L not ordering repression in the baseline model.
3. A post-authoritarian democratic regime embarking on prosecutions of former autocrats.

This state corresponds to one of the baseline model's terminal nodes. In the Markov

extension, repressive agents will have additional actions.

4. A democratic regime following a suppressed potential coup. This state is new and does not have a corresponding node in the baseline game. We assume that should a coup attempt fail, members and leaders of the authoritarian repressive apparatus are permanently removed from office and no longer pose a threat.¹⁰

Figure 1 summarizes our Markov extension.

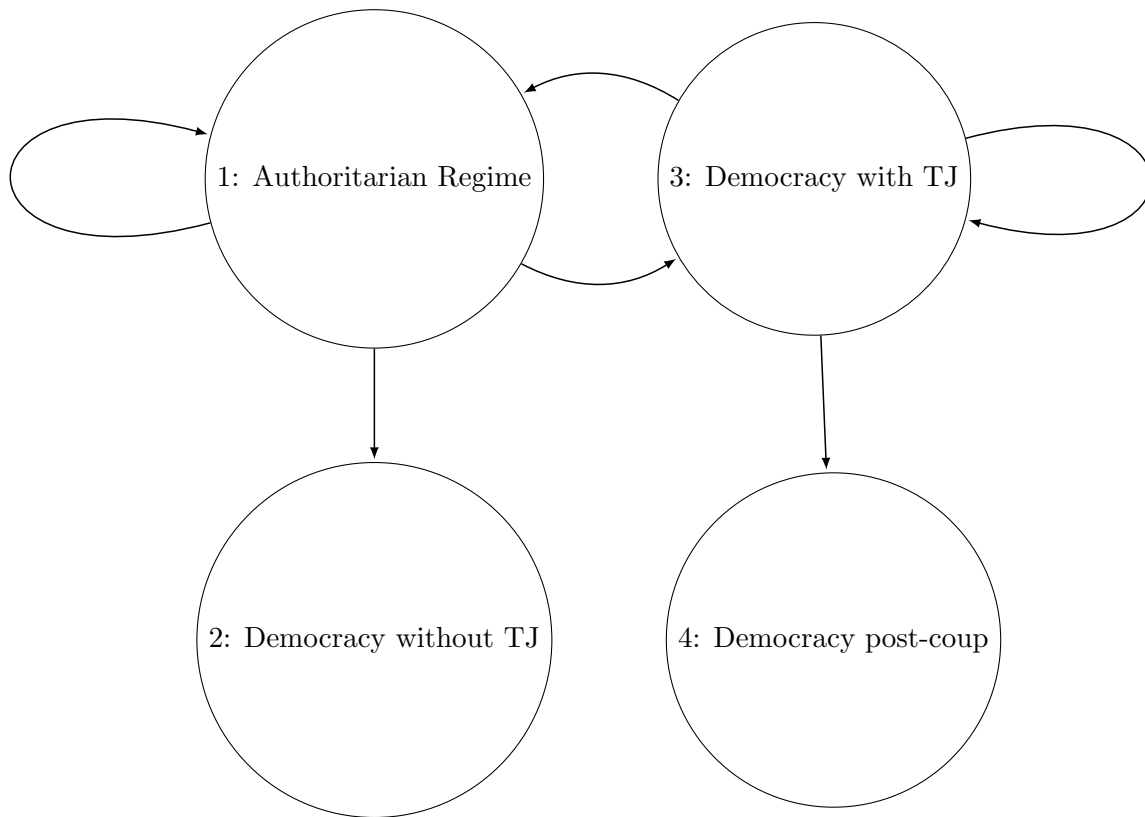


Figure 1: This figure demonstrates the states as defined in the extension.

¹⁰This fourth state is needed to provide a disincentive against coup attempts. If failed coups continued the status quo, there would be no downside in attempting a coup.

We have two parameters. The first is $\gamma \in (0, 1)$, a measure of military strength. γ is used both to define military strength and the probability of a successful coup. The second parameter is δ , the discount factor. We normalize all other payoffs to 0, $\frac{1}{2}$, and 1.

Following TJ prosecutions, if the former autocrats stage a coup, it is successful with some positive probability and the state of the world reverts back to authoritarian rule, at which point the game starts over, in Markov transition fashion.

Strategies, Transition Probabilities, and Payoffs

In state 1, the authoritarian leaders can opt to give or not give an order to repress the citizen population. In the case of an order (and following compensation offer c), the rank-and-file agent of repression then has the choice to carry or not carry out this order, just as in the baseline model.¹¹ State 2 is an absorbing state with no further action (i.e. it transitions to itself with probability 1). In state 3, the post-authoritarian elite select ε which again determines the extent to which the military apparatus will face transitional justice.¹² ε was fixed in the baseline model, but now is a strategic choice made by E . Next, the leaders and their rank-and-file agents of repression have a choice of whether to attempt to stage a coup. State 4 is another absorbing state with no further action.

If an order is given and carried out, transition back to state 1 occurs with probability γ and transition to state 3 occurs with probability $1 - \gamma$. If an order is not given (or not carried out, though this does not happen in equilibrium), transition to state 2 occurs with probability 1. Because state 2 is an absorbing state, the game essentially ends with a transition to state 2. Absent a coup attempt, state 3 transitions to itself with probability 1. If a coup is attempted, transition back to state 1 occurs with probability γ and transition to state 4 occurs with probability $1 - \gamma$. State 4 is again an absorbing state: the game ends

¹¹it is again clear that a leader will never give an order that isn't carried out, and the military will be made indifferent in any equilibrium

¹²We assume, as before, that the order-givers are punished with certainty for the sake of model simplicity. The decision calculus (surrounding coup attempts) for the agents of repression would not be affected were we to allow for flexibility in the leader's punishment as well,

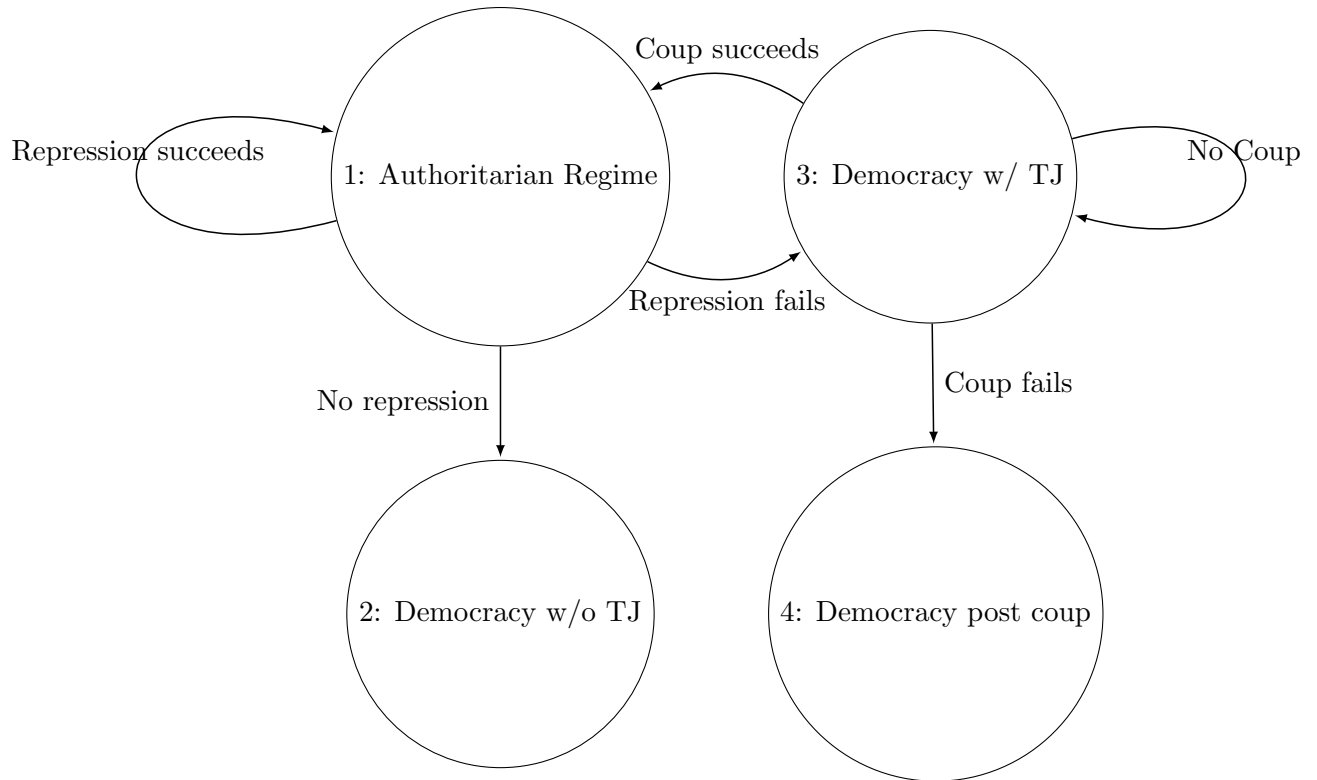


Figure 2: Markov transitions between four states

with a transition to state 4. Figure 2 below summarizes the states and transitions in this Markov extension.

We normalize all payoffs to best encapsulate the present trade-off while facilitating simple analysis. We additionally assume a discount factor of δ .

In State $k = 1$, the authoritarian regime, payoffs in the case of repression are as follows

$$u_L(1, \text{repression}) = 1$$

$$u_R(1, \text{repression}) = c$$

$$u_E(1, \text{repression}) = -1$$

The democratic elite does not take an action in the first period, so we can omit their actions here. When there is no repression, the payoffs are:

$$u_L(1, \text{no repression}) = 0$$

$$u_R(1, \text{no repression}) = 0$$

$$u_E(1, \text{no repression}) = 1$$

We assume that state $k = 2$, Democracy without TJ, has no additional payoffs. The democratic elite has already gained its utility from peace, and there is no need for transitional justice without repression.

The payoffs associated with state $k = 3$, Democracy with prosecutions of the rank and file, are as follows:

$$u_L(3, \text{coup}, \cdot) = 0$$

$$u_R(3, \text{coup}, \cdot) = 0$$

$$u_E(3, \text{coup}, \cdot) = 0$$

The payoffs in state $k = 3$ remain at zero until the next transition to a different state. Note that a coup alone does not affect the democratic elite but, that the coup may cause the state to revert to afterwards that will give the democratic elite a utility of -1.

Absent a coup, the payoffs in state $k = 3$, are:

$$u_L(3, \text{no coup}, \varepsilon) = -1.$$

$$u_R(3, \text{no coup}, \varepsilon) = -\varepsilon.^{13}$$

¹³The payoff of $-\varepsilon$ can be viewed either as a severity of transitional justice or an expected value where severity is 1 with probability ε and 0 otherwise

$$u_E = (3, \text{no coup}, \varepsilon) = 1 + \varepsilon^2.$$

Finally, state 4, the result of a failed coup, hurts all parties, yet maximally hurts the authoritarian leader and members of his repressive apparatus. The payoffs are described by:

$$u_L(4) = -1$$

$$u_R(4) = -1$$

$$u_E(4) = -\frac{1}{2}$$

Note, that no actions are taken in this self-absorbing state and so the payoffs are just a function of the state.

4.1 Results and Discussion

A preview of the results can be characterized as follows: severe transitional justice prosecutions against the rank-and-file agents of repression make democracy less appealing to these agents and incentivize them to join forces with their former leaders in staging a coup d'état. The joining in the coup of the rank-and-file, in turn, increases the chances of its success and subsequent relapse into authoritarianism.

The first effect—the decreased appeal of democracy resulting from transitional justice—comports with the results from the basic game. The second effect—the increased chances of a successful coup resulting from increased prosecutions—is the distinct contribution of the Markov game. Once the possibility of a coup is added, transitional justice prosecutions urge agents of repression to reject the democratic order (which thanks to transitional justice has become more harmful to the military) and push them into supporting a coup. Our results thus speak directly to the question of how to balance using transitional justice prosecutions against the rank-and-file to deter them from obeying cruel orders and provoking them into upsetting the democratic order once they realize that not only their leaders, but also they

themselves are going to be punished.

To solve this game, we start with finding the equilibrium actions in state 3, and then proceed—backward induction fashion—to find the equilibrium actions in state 1. At each stage, we find the optimal actions using Bellman equations and searching for profitable deviations, as is customary in the case of Markov games.

In order to solve for the optimal decision of the democratic elite in state 3, we must first describe the rank-and-file’s choice of whether to stage a coup after transitional justice is implemented. To do so, we introduce some additional, Markov-game-specific notation. Let $V_i(k, (a_i, a_{-i}))$ denote the utility of the player i in state k , given i ’s choice of a_i and other players’ action profile a_{-i} . Next, we can set up the Bellman equation¹⁴ describing the utility of state 3:

$$V_R(3, (coup, \varepsilon)) = \delta\gamma V(1) + \delta(1 - \gamma)V_R(4).$$

The above expression is the expected payoff to the agent of repression from attempting a coup, discounted as it would be consumed in the next stage;

The expression below is the payoff from refraining from a coup and suffering ε of transitional justice and remaining in the democratic state for the next period.

$$V_R(3, (coup, \varepsilon)) = -\varepsilon + \delta V_R(3). \tag{4}$$

Given, the democratic elite’s choice of ε , the agent of repression chooses a coup in state 3, when

$$V_R(3, (coup, \varepsilon)) \geq V_R(3, (nocoup, \varepsilon)) \tag{5}$$

Substituting for $V_R(3, (coup, \varepsilon))$ and $V_R(3, (no coup, \varepsilon))$ using expressions 4 and 5 in the

¹⁴The Bellman equation is simply the expression of $V_i(k, (a_i, a_{-i}))$, given i plays a_i and other players’ strategy a_{-i} in the current period and the discounted value of the states that k may transition into resulting from equilibrium action choice. Since in our game only one player moves per state, it only includes the optimal choice of that player.

Markov framework, we say that the decision to attempt a coup has no profitable deviation when:

$$\delta\gamma V_R(1,.) + (1 - \gamma)(-1)\frac{\delta}{1 - \delta} \geq -\varepsilon\frac{1}{1 - \delta} \quad (6)$$

When $\varepsilon = 0$, which can be interpreted as transitional justice that stops short of reaching the rank-and-file, the condition becomes:

$$\delta\gamma V_R(1,.) + (1 - \gamma)(-1)\frac{\delta}{1 - \delta} \geq 0 \quad (7)$$

The difference between the right-hand-side content of equations 6 and 7 can be interpreted as the difference in the payoff to the rank-and-file under transitional justice that reaches them and transitional justice that only reaches the order-givers.

One sees immediately that the condition expressed in 7 is easier to satisfy than 6, leading to our second lemma.

Lemma 1 *The less transitional justice the rank-and-file agents of repression face, the less likely they are to join a coup. These incentives are minimized when $\varepsilon = 0$, that is, when transitional justice is concentrated on the order givers.*

In state 1, the rank-and-file choose between accepting the compensation for repression or rejecting the order. Given the transition function, γ , this choice boils down to a comparison between transitioning to state 1, the authoritarian regime, with probability γ and transitioning to state 3, democracy post-coup with probability $1 - \gamma$ versus transitioning to state 2 with certainty. That choice can be expressed using the following Bellman equation:

$$V_R(1, (c, \text{accept})) = \gamma c + \delta(1 - \gamma)V(2) \quad (8)$$

$$V_R(1, (c, \text{reject})) = 0 \quad (9)$$

Recall the critical observation from the basic game is that the authoritarian leaders always set compensation level c to make the rank-and-file indifferent between accepting a rejecting. This means that the payoffs expressed in 8 and 9 must be identical leading directly to the lemma below:

Lemma 2 *The payoff to the rank-and-file in state 1, the authoritarian regime, is 0*

We can now use the two lemmas to derive our Markov Perfect Equilibrium. Substituting 0 (based on Lemma 1) for $V_R(1,)$ in expression 6, we have:

$$0 + (1 - \gamma)(-1)\frac{\delta}{1 - \delta} \geq -\varepsilon\frac{1}{1 - \delta}$$

Similarly, substituting 0 for $V_R(1,)$ in expression 7 we have:

$$(1 - \gamma)(-1)\frac{\delta}{1 - \delta} \geq 0$$

or simply

$$0 + \gamma - 1 \geq 0$$

These conditions simplify to $\gamma \geq 1 - \frac{\varepsilon}{\delta}$ and $\gamma \geq 1$, respectively. Notice that when $\varepsilon = 0$, a coup, generically, never occurs. On the other hand, if $\varepsilon > 0$ (i.e. the rank-and-file face transitional justice), there is a very real possibility of a coup, and that possibility is increasing in ε .

Consider now the calculus of the democratic elite. Absent a coup, they find themselves in a peaceful democratic state with payoff $1 + \varepsilon^2$. Under a coup, the payoffs will vary. An unsuccessful coup will be only slightly negative ($-\frac{1}{2}$), whereas a successful coup will launch an authoritarian regime (where, as we show above, the rank-and-file choose repression). Since under the authoritarian regime, the elite payoff is -1 , the elite will choose a value of ε that prevents the rank-and-file from supporting a coup and maximizes its payoff under democracy. The value of ε that satisfies this condition is given by

$$\varepsilon^* = \delta(1 - \gamma) \tag{10}$$

It is easy to see that the democratic elite has no profitable deviations from ε^* : increasing ε would push the rank and file to a coup and their utility would only decrease from decreasing ε .

Given ε^* , we can solve for c^* , the optimal compensation level to be offered to the rank and file in stage 1, the authoritarian regime. Knowing that the leader must ensure indifference for the rank-and-file between accepting and rejecting the repressive order, c^* must satisfy:

$$\gamma(c^* + \delta V_R(1)) + (1 - \gamma) \left(\frac{\delta}{1 - \delta} (-\varepsilon^*) \right) = 0 \tag{11}$$

Note that we use here the fact that ε^* prevents the rank-and-file from staging a coup. Hence, the payoff to the rank and file from state 3, democracy with transitional justice, is simply receiving $-\varepsilon^*$ every turn in perpetuity. Substituting $V_R(1) = 0$ into equation 11 allows us to characterize c^* as:

$$c^* = \frac{(1 - \gamma)^2 \delta^2}{\gamma(1 - \delta)} \tag{12}$$

The above derivation guarantees the repressive apparatus has no profitable deviations from its acceptance range. Equation 10 ensures that that the repressive apparatus similarly has no profitable deviations from “no coup”.

Lastly, we must examine the existence of a repressive equilibrium. This requires analyzing the optimal choice in state 1 and asking if there is a profitable deviation for the leader from ordering repression.¹⁵ For $\gamma = 1$, $c^* = 0$, so the leader will order repression and the equilibrium exists. For $\gamma = 0$, the leader will never order repression. By the intermediate value theorem (and considering that leader utility of repression is monotonic in γ), there exists a γ' such that for all $\gamma > \gamma'$, this repressive strategy profile is in Markov Perfect

¹⁵Recall, that if repression is ordered, the accompanying compensation scheme c^* ensures that it will be accepted.

Equilibrium.

The formal definition of the Markov Perfect Equilibrium has been relegated to the Formal Appendix. Informally it can be summarized in the following proposition:

Proposition 2 *There exists a γ' such that for all $\gamma > \gamma'$ a repressive Perfect Markov Equilibrium equilibrium where*

- *The authoritarian leader offers the rank-and-file a payment $c^* = \frac{1-\gamma)^2 \delta^2}{(1-\delta)^* \gamma}$ in return for repressing anti-authoritarian dissent. The rank-and-file accept.*
- *If repression succeeds, the authoritarian regime continues; if it fails, the democratic elite punishes the rank-and-file with transitional justice $\varepsilon^* = \delta(1 - \gamma)$.*

On part of the democratic elite, this result reflects a balancing act between deterrence and preventing further relapse into authoritarianism. Several additional features characterize this equilibrium. First, note that the agents of repression are compensated for repression in the authoritarian state and are only willing to stage a coup to avoid transitional justice, not out of some sense of loyalty to their former leader. This allows our model to abstract from military cohesion and trade-offs between loyalty and professionalism that have been the focus of some of the literature on non-democratic regimes and turnovers. Second, no matter what kind of pressures from public opinion the democratic elite faces, concerns of avoiding authoritarian backlash will always moderate the extent of transitional justice rank-and file will face.

The level of transitional justice that is tolerated by the rank-and-file decreases, in turn, as the probability that the coup succeeds, γ increases. In the section below, we operationalize this probability of coup success as military strength, or more specifically, military spending.

5 Empirical Implications and Analysis

The goal of the model above was to reconstruct the decision calculus of agents of repression in an authoritarian and post-authoritarian regime facing the prospect of transitional justice.

It systematically and rigorously presented the trade-off facing new democratic elites who on the one hand, wish to deter rank-and-file agents of repression from executing cruel and inhumane orders, but, on the other hand, want to avoid authoritarian backlash. It is safe to assume that truly authoritarian elites—elites that cannot reinvent themselves to flourish in democratic conditions (Grzymala-Busse, 2002)—prefer an authoritarian reversal. But without the participation of the rank-and-file, their attempt to overthrow democracy will not succeed. The main conclusion from our model can be summarized as follows: When the post-authoritarian democratic elite embarks on extensive prosecutions against the former rank-and-file agents of repression, it may endanger the stability of democracy itself.

5.1 Empirical Predictions

This theory generates two key testable, empirical predictions. First, the model suggests that as transitional justice targeting the military apparatus increases in severity, the incentive to stage a coup rises. This leads us to the following hypothesis:

Hypothesis 1 *As the severity of rank-and-file transitional justice increases chances of a successful military coup increase also.*

Second, the model indicates that as military strength increases, the probability of successful coup rises (thus also increasing the incentive to stage a coup). We expect that that:

Hypothesis 2 *An increase in military strength is associate with 1) more coup attempts and 2) a higher percentage of those attempts being successful.*

We will put these hypotheses to the test empirically.

5.2 Research Design

To test the hypotheses of our model, we use data on coups as well as newly collected data on transitional justice prosecutions. Our first dependent variables measure different coup

outcomes (successful or failed coups). In order to test Hypothesis 1, we compare prosecutorial strategies' effects on the occurrence of successful and failed coups. Here, we use successful coups to proxy for coups that have the support of the rank-and-file of the former authoritarian military and security apparatus; failed coups, in contrast, do not have such support. We use Powell and Thyne (2011) 's dataset on coup attempts. These authors code the number of successful (or failed) coup attempts in country i in year t . Using their data, we created a measure for successful coups that takes the value of 0 if either (1) no coup was attempted or (2) a coup was attempted but did not succeed. Since the data in our analysis only include instances with at most one successful coup per country-year, we dichotomize this variable. The variable "Successful" takes the value of 1 when a coup was attempted and was successful, while 0 represents a country year with no coup attempt or no successful coup attempts. In the case of failures, there are a few instances with two failures in the same year. Hence, the variable "Failed" is a count and takes on the number of coup failures in a given country in a given year. It assumes 0 when there were no attempts or all attempts succeeded.

For transitional justice prosecutions, we use the Transitional Justice and Democratic Stability Lab's recently-released Global Transitional Justice Dataset (GTJD) (Bates, Cinar and Nalepa, 2020), which now covers data on criminal prosecution events related to former authoritarian elites. The GTJD contains data on all countries which have transitioned to democracy since the end of World War II.¹⁶ Events related to criminal prosecutions include indictments, arrests, trials, convictions, amnesties, appeals, annulments, and acquittals, as well as information about the creation of institutions for conducting criminal trial of former perpetrators of human rights, such as special courts. Members of the TJ and Democratic Stability Lab collected data on these events beginning with each country's democratic transition¹⁷ and continuing through 2018 or the year in which the country in question ceased to be democratic.¹⁸

¹⁶If a country has transitioned multiple times, the GTJD uses its most recent transition.

¹⁷A democratic transition is coded according to Geddes, Wright and Frantz (2014) or as following the first fully democratic national election.

¹⁸It is worth noting that this conservative demarcation of the democratic spell actually biases our results

The criminal prosecution events were then coded as “positive”—or moving the criminal trial process forward—or “negative”—or delaying the trial process or even reversing it.¹⁹ The prosecution events were also coded to account for the targeted perpetrators: an event is coded as targeting leaders, the rank-and-file, both, or neither. The total positive and negative events of each type were then compiled into a time series panel of country-years.

Using this data from the GTJD, we constructed a measure of annual prosecution severity for each type of targeting:

$$Severity_{i,t} = \frac{\sum_1^{t-1} P_{it}}{\sum_1^{t-1} (P_{it} + N_{it} > 0 + 1)}.$$

The 1 is added to the denominator in order to keep the measure defined.²⁰

After calculating severity measures for each type of target (leadership or rank-and-file), we then constructed a measure of annual *relative severity*. This measure is equal to the rank-and-file severity of a given year, divided by the rank-and-file severity + the leadership severity + 1. Again, the 1 in the denominator keeps the measure defined:

$$\text{Relative Severity}_{i,t} = \frac{Severity_{i,t}^{RF}}{Severity_{i,t}^{RF} + Severity_{i,t}^L}.$$

In order to test Hypothesis 2—that military strength increases the probability of both coup attempts and successful coups—we include data on military strength in all of our analyses. We use military expenditures as a percentage of GDP because it is likely that stronger militaries demand larger percentages of state resources.

For control variables, we included a dummy variable for the Cold War, where years against our hypotheses because coups occurring at the very early period of the end of the authoritarian regime, before elections have taken place, are omitted from our analysis. In other words, cases where the military perceives the actual or incoming targeting of the rank-and-file to be happening and consequently launches a coup while still in the transitional period are not included in our dataset.

¹⁹Note that we refrain from identifying trial events as positive or negative in the normative sense. So while an acquittal due to procedural errors may be seen as positive because it is a sign that the justice system is working fairly, we nevertheless code it as “negative.”

²⁰This is in line with the measure of severity from (Bates, Cinar and Nalepa, 2020), though a key difference is that our severity measure is calculated annually.

earlier than 1990 are coded as 1, and years 1990 and after as 0. Controlling for Cold War dynamics—either in the form of direct support from or the potential of intervention from the United States, the Soviet Union, or one of their powerful allies with an interest in provoking a coup—allows us to concentrate better on testing the calculus of the rank and file. We also control for economic conditions that might affect the opportunity costs of coup supporters. For instance, even if the former rank-and-file will not be punished, if they cannot find gainful employment, they may become more likely to support a coup d'état. In order to account for this possibility, we also included yearly GDP per capita (logged) data from the World Bank.

Using this panel data, we use OLS to fit a series of linear probability models. We include country and year fixed effects in all regressions in order to better approximate causal interpretation.

5.3 Empirical Results

Table 1 provides the results of our first analysis, testing hypothesis 1. In the first model, the dependent variable is successful coups, while the second model provides results using failed coups as the dependent variable. As we can see, the relative severity of targeting of rank-and-file perpetrators is a positive and statistically significant predictor of successful coup attempts, while other variables like the Cold War, GDP per capita, and military spending are not. That is, the targeting of the rank-and-file is positively associated with the onset of successful coups.

We also sought to compare these results with a regression of the occurrence of failed coups, i.e. coups that were less likely to have the support of the rank-and-file. The results are presented in the second model of Table 1. Figure 3 illustrates the results of this comparison directly. We see that while relative severity is statistically significant and has a positive coefficient for successful coups, it is not statistically significant for failed coups. Moreover, the sign on the coefficient for failed coups is negative. Generally, this provides suggestive

Table 1: Effect of Rank-and File Criminal Prosecution on the Probability of a Successful or Failed Coup Attempt

	Successful	Failed
Relative Severity	0.115* (0.048)	-0.005 (0.089)
Cold War	-0.007 (0.012)	0.030 (0.022)
GDP Per Capita	-0.000 (0.000)	-0.000 (0.000)
GDP military spending	0.159 (0.518)	-1.159 (0.965)
Constant	-0.002 (0.012)	0.041 (0.021)
Observations	918	918
Adjusted R^2	-0.047	-0.050
BIC	-2114.709	-973.318
Degrees of Freedom	870.000	870.000

Standard errors in parentheses

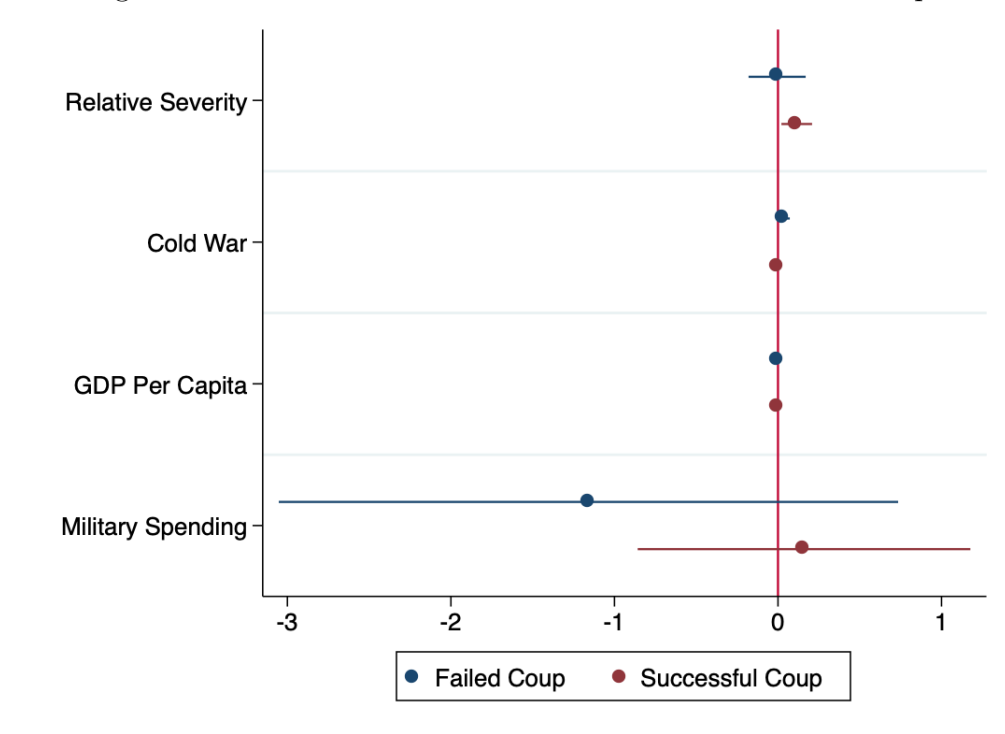
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

evidence in support of our argument. Targeting the rank-and-file, which we capture using our measure of relative severity, leads to an increase in the likelihood that a coup attempt will succeed, but does not increase the likelihood of those coup attempts that fail.

As shown in Table 1 and Figure 3, other variables are not statistically significant. Importantly, this includes our measure of military strength: a higher percent of GDP spent on the military does not predict an increase in the likelihood of a coup attempt, successful or otherwise. This lack of significance could be attributed to the poor quality of our measure of military strength. While powerful militaries often demand a greater share of resources, a measure as blunt as the percent of GDP spent on the military may capture other dynamics that have little to do with military strength.

Finally, we provide results for an alternative hypothesis: that countries whose regimes were military dictatorships are more likely to have a successful coup, since the military is the

Figure 3: Coefficient Estimates for Failed vs. Successful Coups



principle coup actor. To test this, we repeat the analysis of relative severity and successful coups—the analysis outlined in the first model of Table 1—but subset the data by the regime type of the former authoritarian regime. All explanatory variables remain the same. Table 2 shows the results of these regressions. The first model presents the results of our analysis for only those countries that transitioned from a military regime. The second model presents the results of our analysis for countries that transitioned from all other authoritarian regime types.

As both models in Table 2 show, once we partition our data by ancien regime type, relative severity (of rank-and-file prosecutions) is not a statistically significant predictor of successful coups. This result can be attributed to the relatively small number of observations in each subset. Nevertheless, this finding does not provide evidence that there is reason to think that the prior regime type exudes a significant influence over the occurrence of successful coups. Furthermore, in line with the results of our prior analysis, military spending as a percent of GDP is also not a significant predictor of a successful coup, even in democracies

Table 2: Effect of Rank-and-File Criminal Prosecution on the Probability of a Successful Coup in Democracies Transitioning from Military vs. Non-military Regimes

	Military	Non-Military
Relative Severity	0.162 (0.091)	0.123 (0.068)
Cold War	-0.003 (0.014)	-0.002 (0.030)
GDP Per Capita	-0.000 (0.000)	-0.000 (0.000)
GDP military spending	-0.529 (1.021)	0.264 (0.720)
Constant	0.013 (0.023)	-0.004 (0.014)
Observations	316	538
Adjusted R^2	-0.036	-0.058
BIC	-776.274	-1118.450
Degrees of Freedom	300.000	504.000

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

that have transitioned from military regimes.

Our empirical analysis thus provides mixed support for the theory we model in Sections 3 and 4. While the targeting of rank-and-file members of the former regime does appear to increase the likelihood of a successful coup, the strength of the military appears to have little effect. This is true regardless of the type of regime that existed prior to the democratic transition.

5.4 Egyptian case study

Egypt's post-authoritarian experience serves as convincing illustration of the dynamics of our model in real world settings because it experienced a successful coup so soon after the transition.

In 2011, near 30-year rule of Hosni Mubarak came to a halt with the Arab Spring, as

Egypt became the second country in the region to experience a revolution on the wave of the Arab Spring.

Shortly after toppling Mubarak, a transitional government was installed which scheduled fully democratic elections for mid-2012. During this interim period preceding elections, the new government instituted a number of transitional justice policies, which largely focused on issues related to the non-rank and file.²¹ In March 2011, a new national security agency replaced the Mubarak-era State Security Investigations agency. Three months later, Mansour el-Eissawy, the new Interior Minister, fired 669 top police officers—major generals, brigadier generals and colonels from the officer corps. Over the course of 2011 and 2012 a number of former cabinet members and regime elites were tried and convicted for corruption and graft related offences that they had allegedly been involved in under Mubarak. Finally, Mubarak himself along with six former high-ranking members of his Ministry of Interior was put on trial for the murder of anti-government protesters. All were convicted in 2012.

Despite these actions, during the period of the transitional military government, the prosecution of the rank-and-file was minimal. In March 2011, for example, a military court acquitted an army doctor accused of carrying out forced virginity tests on female prisoners, despite the Minister of Defense having admitted that the incident happened. The election of Mohamed Morsi in June 2012 marks the beginning of our coding for the democratic period. In August 2012, Morsi, in an effort to reign in the influence of the military, forced the retirement of a number of its senior leadership, including chairman of the Supreme Council of the Armed Forces Mohamed Hussein Tantawi. It was under the Morsi regime that prosecutions and convictions of rank-and-file agents of repression began in earnest. For instance, in March 2013, a police sniper was convicted for shooting at protestors during the revolution, including intentionally aiming at their eyes, and sentenced to 5 years in prison.

Only four months after this conviction, in July 2013, the military launched a coup that overthrew Morsi's government and established a new regime lead by General Abdel Fattah

²¹Note that due to the conservative demarcation of the democratic period in the Global Transitional Justice Dataset, none of these pre-electoral TJ policies are reflected in our analysis.

el-Sisi. This coup was ostensibly motivated by a variety of factors. Reporting indicates that military leadership was concerned about Morsi's rhetoric implying foreign escapades in Sudan, Ethiopia and Syria. Morsi's purge of military leadership and his threats to the military's economic power were also reported to motivate coup leaders, while, recent protests against Morsi's government provided the opportunity to launch the coup itself.

Although prosecutions of the rank-and-file have not been explicitly mentioned as a reason behind the coup, reporting is generally centered on the motivations of those launching the coup, not those who chose to support it, such as the rank-and-file. Moreover, coup launchers are more likely to justify their military takeovers with arguments that focus on ideals of law and order rather than personal security and maintaining their own positions and power. The Egyptian coup leaders focused their justifications on preventing foreign entanglements.

To summarize, the new transitional government that succeeded Mubarak in 2011 pursued a number of transitional justice policies and prosecutions largely ignoring the rank-and-file. The newly elected Morsi government from 2012-2013, in contrast, initiated purges of the military leadership, and began the prosecution and conviction of rank-and-file perpetrators. Within four months of the first rank-and-file conviction, the military launched a successful coup, restoring military government.

6 Conclusion

Societies transitioning from authoritarianism and civil war to a democratic order have to reckon with members and leaders of their repressive apparatus. Former agents of repression can either be integrated back into the new democratic state, following extensive vetting (Nalepa, 2010; Stan, 2013) or they can be prosecuted.²² Yet even the decision to prosecute is not a straightforward one. In this paper, we have focused on where new democracies should strike the balance between prosecuting leaders of the repressive apparatus and prosecuting

²²Technically, they can also be purged and forced to seek employment elsewhere, but this strategy comes with other risks, such as increased levels of organized crime.(Bates et al., 2020; Lessing, 2017)

the rank-and-file members of that apparatus.

We have proposed a Markov Transition game that simultaneously captures the decision calculus of democratic elites resolving where to strike that balance and the former repressive apparatus members deciding whether to use repression and, later, whether to stage a coup d'état. Our model's parameters, hence, correspond to key arguments that have been made in support of wide-spread prosecutions of former perpetrators of human rights, as well as the key arguments made against of such prosecution.

The reason to prosecute agents of repression is simple: without sanctions for carrying out the most inhuman orders, agents of repression unscathed, will remain recruitable by *future* autocrats. Retroactive justice, no matter how legally complicated it is to pull off, acts as a deterrent for future autocrats.

The counterargument to prosecutions focuses on threats to democratic stability. With nothing to lose but their freedom, what stops the agents of repression from casting their lot with the former authoritarian leaders and helping them stage a coup? Democracy with transitional justice may just tip the balance for rank-and-file agents of repression, to make them eager recruits of their *former* bosses.

In the solution to our model, we find that there is a critical level of military strength at which a repressive equilibrium is possible. In this equilibrium, autocrats mobilize their agents of repression to crack down on democratic resistance. When despite the crackdown, democracy eventually succeeds, democratic elites must pull their punches to avoid an authoritarian reversal. Were they to embark on wide scale prosecutions, extending beyond the top echelons of power, they just might upset the democratic order they so hard fought to establish.

This finding microfounds theoretically what many scholars of democratic transitions, from Huntington (1993) to Vinjamuri and Snyder (2004) have argued all along: that harsh transitional justice plays into the hands of spoilers of democracy and rule of law. At a practical level, it also explains why justice NGOs and INGOS will always remain dissatisfied

with the level of criminal transitional justice that is observed around the world. While we share the sentiment behind their lamenting of impunity for those who “killed with their own hands” [p. 114](Todd et al., 2000*a*) in the name of the dictatorship they worked for, to insist on harsher prosecutions may threaten democratic stability. This doesn’t mean we have to throw the baby out with the bathwater and eschew all forms of transitional justice. Non-criminal forms of TJ, such as truth commissions and lustration, have proven to be not only a valuable, but in some instances, an indispensable form of reckoning with the past Ang and Nalepa (2019).

References

- Ang, Milena and Monika Nalepa. 2019. "Can Transitional Justice Improve the Quality of Representation in New Democracies?" *World Politics* 71.
- Bates, Genevieve, Ipek Cinar and Monika Nalepa. 2020. "Accountability by Numbers: Evidence from a Global Transitional Justice Events Dataset (1946-2016)." *Perspectives on Politics* 18:161–184.
- Bates, Genevieve, Ipek Cinar, Monika Nalepa and Evgenia Olimpieva. 2020. "What is the effect of Personnel Transitional Justice on Crime?".
- Brumwell, Stephen. 2018. *Turncoat: Benedict Arnold and the Crisis of American Liberty*. Yale University Press.
- Capoccia, Giovanni and Grigore Pop-Eleches. 2020. "Democracy and Retribution: Transitional Justice and Regime Support in Postwar West Germany." *Comparative Political Studies* 53(3-4):399–433.
- Dragu, Tiberiu and Yonatan Lupu. 2018. "Collective action and constraints on repression at the endgame." *Comparative Political Studies* 51(8):1042–1073.
- Elkins, Stanley and Eric McKittrick. 1995. *The Age of Federalism: The Early American Republic, 1788-1800*. Oxford University Press.
- Elster, Jon. 2004. *Closing the books: Transitional justice in historical perspective*. Cambridge University Press.
- Geddes, Barbara, Joseph Wright and Erica Frantz. 2014. "Autocratic breakdown and regime transitions: A new data set." *Perspectives on Politics* 12(2):313–331.
- Gehlbach, Scott. 2013. *Formal models of domestic politics*. Cambridge University Press.
- Grzymala-Busse, Anna M. 2002. *Redeeming the communist past: The regeneration of communist parties in East Central Europe*. Cambridge University Press.
- Human Rights Watch. 2001. "Reports: Argentina."
- Human Rights Watch. 2005. "Argentina: Amnesty Laws Struck Down."
- Huntington, Samuel P. 1993. *The third wave: Democratization in the late twentieth century*. Vol. 4 University of Oklahoma press.
- Kaminski, Marek M and Monika Nalepa. 2014. "A model of strategic preemption: why do Post-Communists hurt themselves?" *Decyzje* 21:31–66.
- Kaminski, Marek Mikolaj. 2019. "Generalized Backward Induction: Justification for a Folk Algorithm." *Games* 10(3):34.

- Krcmaric, Daniel. 2018. “Should I stay or should I go? Leaders, exile, and the dilemmas of international justice.” *American Journal of Political Science* 62(2):486–498.
- Lessing, Benjamin. 2017. “Counterproductive punishment: How prison gangs undermine state authority.” *Rationality and Society* 29(3):257–297.
- Masri, Safwan M. 2017. *Tunisia: an Arab anomaly*. Columbia University Press.
- Myerson, Roger B. 2008. “The autocrat’s credibility problem and foundations of the constitutional state.” *American Political Science Review* 102(1):125–139.
- Nalepa, Monika. 2010. “Captured commitments: an analytic narrative of transitions with transitional justice.” *World Politics* 62(2):341–380.
- Nalepa, Monika and Emilia Justyna Powell. 2016. “The role of domestic opposition and international justice regimes in peaceful transitions of power.” *Journal of Conflict Resolution* 60(7):1191–1218.
- Nino, Carlos Santiago. 1998. *Radical evil on trial*. Yale University Press.
- Powell, Jonathan M and Clayton L Thyne. 2011. “Global instances of coups from 1950 to 2010: A new dataset.” *Journal of Peace Research* 48(2):249–259.
- Stan, Lavinia. 2013. “Reckoning with the Communist past in Romania: A scorecard.” *Europe-Asia Studies* 65(1):127–146.
- Todd, Stephen Charles et al. 2000a. *Lysias: Against Erasthenes*. Vol. 2 University of Texas Press.
- Todd, Stephen Charles et al. 2000b. *Lysias: Against Manthitheus*. Vol. 2 University of Texas Press.
- Tyson, Scott A. 2016. “The agency problem underlying the use of repression.”
- Vinjamuri, Leslie and Jack Snyder. 2004. “Advocacy and scholarship in the study of international war crime tribunals and transitional justice.” *Annual Review of Political Science* 7:345–362.

A Formal Appendix

Here we describe the full repression equilibrium for the Markov extension. We claim that for sufficiently high γ , our parameter of military strength, the strategy profile described below is in Markov Perfect Equilibrium:

Let $c^* = \frac{(1-\gamma)^2\delta^2}{\gamma(1-\delta)}$ and $\epsilon^* = \delta(1 - \gamma)$. The repressive Markov Perfect Equilibrium strategy profile is:

State 1: $(c^*, 1; [c^*, \infty);)$

State 2: no actions

State 3: (\cdot ; 0; ϵ^*)

State 4: no actions

B Empirical Appendix

Figure 4: Coefficient Estimates for Successful Coups

