

# Clean sweep or picking out the ‘bad apples’: the logic of secret police purges with evidence from Post-Communist Poland

Monika Nalepa<sup>1</sup> and Barbara Piotrowska<sup>2</sup>

<sup>1</sup>Department of Political Science, The University of Chicago, Chicago, IL 60615, USA

<sup>2</sup>Department of Political Economy, King’s College London, London, UK

May 4, 2022

## Abstract

Peaceful regime transitions present new democracies with the dilemma of how to reckon with members of their enforcement apparatuses. One solution is to disband authoritarian agencies and build new ones. A competing approach, is to fire only the “bad apples.” Strapped for competent agents, new democracies may take the latter approach, particularly if the preceding autocratic regime employed predominantly competent agents. We use a formal model to compare conducting 1) a “thorough purge”— dismissing everyone —2) a “selective purge,” where agents are evaluated on a case-by-case basis, and 3) retaining everyone. We predict that competence of agents of the prior regime moderates how their ideological extremity favors selective purges over thorough or no purges. We corroborate our theory using a paired comparison approach, employing new data from the Institute of National Remembrance (IPN) in Poland on the operation of verification commissions in 49 sub-national regions in 1990.

Acknowledgments: The authors wish to thank Genevieve Bates, Olga Gasparyan, Scott Gehlbach, Hanzhang Liu, Zhaotian Luo, Molly Offer-Westort and Hongding Zhu for comments on earlier versions of this draft.

# 1 Introduction: Purge or reform?

When in 2014, the world turned its attention to Euromaidan in Kiiv Ukraine, it was in response to the brutal pacification of peaceful protesters in the Central square of Kiiv, the Maidan. People had gathered there in opposition to the ruling government's pulling out of a cooperation agreement with the EU in what was perceived as a pro-Russia move.

The government led by Victor Yanukovich was in the last months of what was generally perceived as his final term in office. After days of trying to deflect the protest, Yanukovich resorted to sending the most brutal troops of riot police, the infamous Berkut unit, to pacify the demonstrators. Police fired on the occupants of Maidan and as a result of their actions, more than a hundred protesters died and many more suffered injuries. Europe watched in shock as a peaceful protest turned bloody when police did not hesitate to obey inhumane orders. To those familiar with Eastern Europe's communist past, it brought back memories of brutal police crackdowns on peaceful anticommunist demonstrations in Poland, Hungary, Czechoslovakia, Romania, and others. Why were the police so eager to follow inhumane orders coming from the pro-Russia executive?

This case illustrates the potential costs of refraining from police reform in the aftermath of regime change. Although Ukraine had fifteen years to reform its enforcement apparatus and turn a militia created to serve Soviet autocrats into a police designed to serve citizens, it did very little in this time span to rid the enforcement apparatus of its "bad apples."<sup>1</sup>

Ukraine was not an isolated case. Following the dissolution of the Soviet Union, its satellite states were faced with the challenging task of purging their security and enforcement apparatus of officers whose responsibilities ranged from spying and repressing the opposition to fighting white-collar crime in communist states. In the words of one Polish historian: "there were two opposing views among the [new] political elite in response to how the secret police treated the anti-communist dissident movement. Some wanted to limit the influence

---

<sup>1</sup> See graph in the Appendix, based on the GTJD indicating how little Ukraine did exactly to reform its enforcement apparatus following the Orange Revolution.

of old agencies entirely, resulting in a thorough purge. On the other end of the spectrum were the pragmatists [who] argued that law enforcement is but a tool in the hands of politicians and ... [t]he secret service should be reformed, but in a way that retains the experts whose origins were in the communist secret police force” (Kozlowski 2019, p.289).

This quote suggests that the personnel choices before newly democratized states seeking to reform their security services were to either purge<sup>2</sup> the enforcement apparatus in its entirety or to try and “pick out the bad apples.” There was also a third possibility of doing nothing.

Our aim is to explain how politicians following regime change settle on reforms of its enforcement apparatus. In contrast to work on authoritarian purges Montagnes and Wolton (2019); Crabtree et al. (2020), we focus on the post-authoritarian period and on decisions pertaining to low level enforcement agents. Nevertheless, the topic of our article is purges in the transitional justice context.

The article builds on a key theoretical insight: We show that the meaning of “bad” in the phrase “picking out the bad apples” is ambiguous. In the context of regime change, it could refer to officers of enforcement who resorted to extreme violence, persecuted the anti-authoritarian opposition and “deserve” to be punished by members of the incoming democratic regime; or it could refer to incompetent agents, who were appointed to their positions for their ideological loyalty to the communist leaders. The latter kind of officers is bad not because they persecuted the opposition, but because they are incompetent and retaining them offers no value to the new state.

Building on this insight, we use the familiar literature on the loyalty-competency trade-off to frame this dilemma of new democracies in a simple formal model. Key in modeling the incentives of politicians in the new democracy is the desire to staff the security and enforcement apparatus with agents who are at once loyal and competent. However, in our

---

<sup>2</sup> We note that although the concept of purges may suggest similarities to ideas of disbanding the police, reminiscent of postulates from the US 2020 Uprising, postulates for “disbanding” or “defunding” the police did not include ideas of replacing it with a new police force. See Bates and Cole (2021).

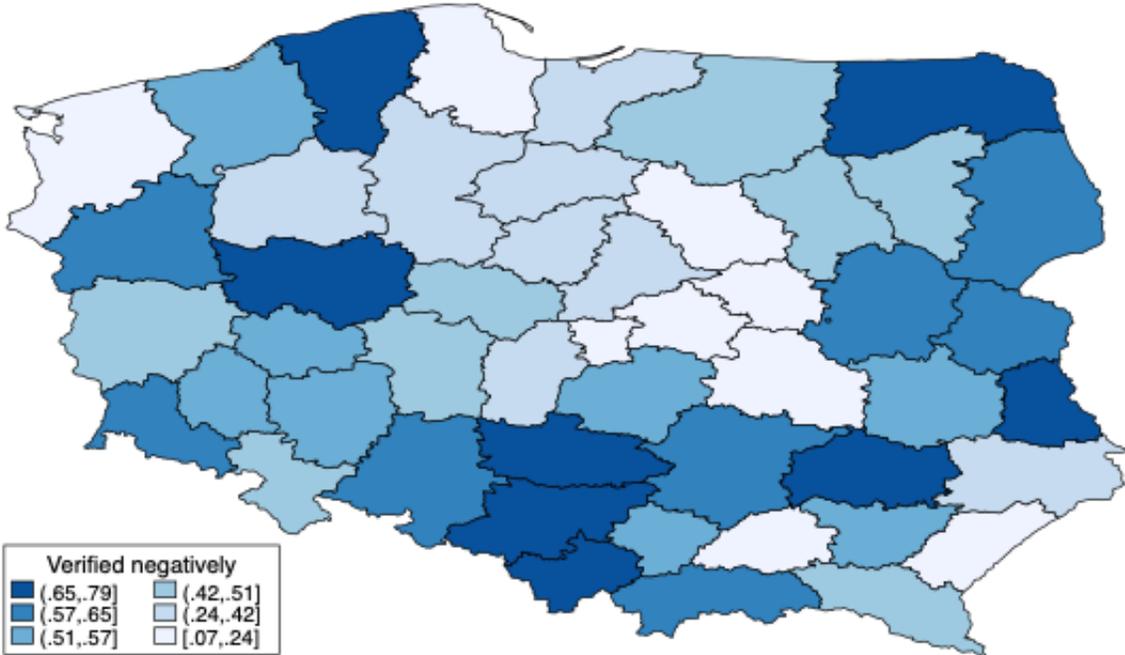
model, rather than being endpoints of a spectrum, loyalty and competence are separate, sometimes even orthogonal dimensions to one another.

This sense of practicality—the demand for experts skilled in fighting crime — we argue, keeps new democratic politicians’ desire for transitional justice (and potentially revenge) in check. Hence, the extent of repression does not affect the type of purge directly but is moderated by agent competence.

To illustrate our theory, we use data from another former socialist state that at the critical juncture of communist collapse faced the challenging task of purging its security and enforcement apparatus: Post-Communist Poland.

Poland is ideal for testing our theory of how policy-makers resolve the dilemma of whom to fire and whom to rehire because it decentralized its so-called verification process, delegating to its 49 regions and 49 regional verification commissions the decision of how to deal with each and every employee of the former secret police. Figure 1 created using data from the Institute of National Remembrance (IPN) reveals a high level of variation in RVC decisions.

**Figure 1:** *Geographic distribution of the proportion of secret service denied employment by the new democratic state*



The proportion of those verified negatively (denied employment by the new democratic state) ranges from 0.07 to 0.79 in the period preceding the appeals process.

The variation in Poland’s verification outcomes makes it ideal for studying our question. It allows us to test the findings from our formal model using the paired comparison technique, a method that has traditionally been used for generating theory (Gerring 2016). Sydney Tarrow (2010), in a first article-length discussion of it in political science, likens it to experimental designs in that it allows the researcher to isolate the impact of a “single variable or mechanism on outcomes of interest” (Tarrow, p. 224). To our knowledge, the method has never been used in combination with formal modeling, even though it is actually ideally suited for testing propositions based on comparative statics of equilibrium results in game theory.<sup>3</sup> The insights generated from comparative statics are often conditional on the keeping values of other variable below or above a certain threshold. This is the case also in our theory, where the impact of competent agents moderates the predicted effect of extreme agents on the probability of taking the time to “pick out the bad apples” (what we call, a selective purge). Thus, our proposition, and ultimately hypotheses translate directly into a paired comparison research design. The fact that all the regional verification commissions in our data come from the same country allows us to control for any variables not in our model, while the parameters that are modeled guide our choice of specific regional commissions to match the specific conditions described in our hypotheses.

## 2 Loyalty, Competency and Justice

The recent decade has brought renewed interest among scholars in how autocrats use their various enforcement agencies to stay in power.

Sheena Greitens (2016) provides a book-length treatment of this dilemma, arguing that the way autocrats organize their repressive apparatuses depends on the kind of threat they

---

<sup>3</sup> In one of the few discussions of combining qualitative evidence with formal models, Lorentzen et al. (2017) note that the overwhelming majority of papers providing such evidence published in top journals and presses use qualitative cases heuristically rather than systematically, as evidence.

face. If they are more concerned with revolution from below, they will centralize secret police forces and create a powerful agency that might be threatening to some autocrats, just not those who predominantly fear a popular rebellion. If their primary concern is with a lateral coup d'état, they will fragment their security forces and sacrifice efficiency in intelligence gathering to avoid a threat to their power.

Poland, the case we use here, was most definitely an instance where the greatest threat to communist rule came in the form of revolution from below rather than a lateral agency. Hence, the secret police was highly centralized. Another important feature of Communist secret police agencies was their surprising competence. In a recent article, Scharpf and Glasel (2020) present a mechanism leading to low competence in the security apparatus of some former authoritarian regimes, such as Argentina. According to their argument, the career incentives in hierarchical organizations, characteristic of authoritarian regimes made it hard for mediocre agents to advance. The arduous character of secret police work (surveillance, long and unpredictable hours) made it a line of work that was attractive for ambitious yet lacking talent candidates.

Our evidence from Post-Communist Europe does not corroborate this story. Based on an analysis of applicants to the communist police forces, Oseka (2008) finds that although entry position applicants did indeed come from rural areas and rarely carried even a high school diploma, this was no longer the case for seasoned employees of secret police who chose the best and brightest of the police and put them through school, college, and rigorous training. According to an internal survey prepared to accompany exit interviews (Komar and Niedzialek 1990), the Polish secret police agency was generally a source of social mobility that turned uneducated, poor and rural unskilled workers into professionals with degrees (at times from college) living in their own apartments in urban areas.

Typically, the best police officers, rather than the worst, would be transferred from the regular citizens' militia to the secret police forces. To merit such transfer, they had to exhibit potential for learning the most valuable skill a secret police officer could acquire: the ability

to recruit informers. This was a complex task and one on which the secret police agencies spent considerable resources, particularly in areas where the battle for citizens' hearts and minds was most heated (Piotrowska 2020).

Indeed, the ability to have either competent or loyal agents working for the autocratic rulers was the inspiration behind formalizing the so-called loyalty competency trade-off in the first place.

Originally modeled by Egorov and Sonin (2011), the loyalty-efficiency trade-off was used to explain why rulers, but particularly dictators, have to balance their demand for loyal agents on the one hand with the need for skilled ones, on the other. In the words of these authors, "the very competence of the vizier makes him more prone to treason" (Egorov and Sonin 2011, p. 904). Unsurprisingly, strong rulers will resolve the dilemma differently than weak rulers. The former will choose competence over loyalty, as they are capable of staging off threats; the latter will invest in loyal agents as it does not take much for them to feel threatened.

Building on Egorov and Sonin (2011) foundational model, Alexei Zakharov (2016) considers interaction between a dictator and individual subordinates where the dictator chooses a specific level of loyalty (and correspondingly lack of competence) he needs from his subordinates. A controversial prediction from his model is that unskilled agents invest in loyalty because they are aware that their limited skill-set renders them unemployable by future rulers. But skilled agents, knowing that they can find lucrative employment under any regime have an incentive to shirk.

The next models we discuss propose ways of alleviating the trade-off between loyalty and competence. Josef Woldense (2018) integrates social networks into the analysis to show that shuffling employees laterally from one geographic region to another can prevent agents' entrenchment and staves off their becoming powerful enough to threaten the ruler. At the same time, shuffling is better than firing, because agents continue to acquire expertise and have higher competence value to the ruler. Shuffling is hence one of the ways to avoid the

loyalty-competence trade-off. It allows rulers to have skilled employees that do not pose a challenge to their rule.

Yet not posing a threat does not immediately imply loyalty. What if the agent refuses to put in the effort? Jack Paine (2022) examines this aspect of the competency-loyalty trade-off in the context of organizing security and repressive apparatuses by authoritarian rulers. In this interpretation of the trade-off, the authoritarian ruler considers appointing a professional (hence equipped with expertise) army against a personal (hence loyal, but relatively less skilled) militia. It follows from the fact that neither kind of security force can survive a revolution from below (the radicalism of such revolutions reduces survival prospects of both kinds of security apparatuses to nil) that autocrats will appoint personalist rather than professional militaries when they are more concerned about external threats than about revolutionary threats from within. Because, in the view of Paine (2022) post-revolutionary regimes fire everyone from the enforcement apparatus, professional militaries lack the will to effectively fight revolutionary threats. In contrast, in the case of external threats, professional militaries can count on leniency. Indeed, in some instances, they even collaborate with external challengers to the autocrat. In light of this, they put less effort into averting external threats. Consequently, the autocrat who fears such threats is better off investing in a personal militia.

Our model, similarly to Paine's and Woldense's, averts the trade-off between loyalty and expertise but in yet a different way: agents with expertise and weaker loyalty to the outgoing, authoritarian regime, may in fact be more loyal to the new democratic regime. At the same time, our model departs from all four discussed above in that we are less interested in the former authoritarian ruler and more in the incentives of democratic rulers who succeed him. It also departs from the broader literature on the loyalty-competence trade-off in that in contrast to this literature it is not concerned with leadership positions within the authoritarian regime. Note that a "competent vizier" poses a threat to the sultan precisely because of his competence. Hence one cannot at once be competent and loyal. It

is not so with our model that focuses on low level agents of repression who could not post a threat and moreover, as we explain below can be competent and loyal at the same time.

The fourth feature of our model is that since the new democrats inherit a state staff potentially loyal to the autocrat, their decision, at least initially, is more about whom to appoint, than whom to purge from the authoritarian state apparatus.

Not modeled in our theory, as a normative motivation cannot be incorporated in a non-cooperative game, is the decision to fire agents based on their repressive behavior towards dissidents. We return to the prevalence of repressive behavior as a motivation for firing security agents as a possible alternative explanation in the empirical section of our paper, devoted to testing our theory via the paired comparisons approach.

### 3 The Model

In order to model the tasks before new democratic politicians deciding on how to deal with members of an authoritarian security apparatus, we follow Nalepa (2022) in using the workhorse model of American politics scholars studying bureaucracies, the delegation model (Epstein and O'halloran 1999; Huber and Shipan 2002; Callander et al. 2008). The problem of whether or not to fire a law enforcement agent is the flip side of the delegation problem. Instead of choosing to delegate, the politician in a new democracy chooses whether to retain an existing agent that was appointed by a different (authoritarian) principal.

In the case of the decision to purge, the new democratic politician can retain an agent of the former authoritarian regime (equivalent to delegating in models of bureaucracy) or can purge this agent and appoint an inexperienced replacement in his place.<sup>4</sup> The players are a (new democratic) Politician with ideal point 0 in a uni-dimensional issue space,  $S$ , and a former law enforcement agent whose type is drawn from two cross-cutting type spaces:

1.  $T^C = \{competent, incompetent\}$

---

<sup>4</sup> We note that while the original agent, as we show below, can also be incompetent/inexperienced, the new one definitely is such.

2.  $T^I = \{moderate, extreme\}$

$T^I$  represents the fact that the agent of the law enforcement apparatus has an ideal point in the policy implementation space, which is unknown to the Politician, but can be either  $x_M$ , if the agent is moderate or  $x_E$ , if the agent is extreme. An agent is moderate with probability  $\theta \in (0, 1)$  and extreme with probability  $1 - \theta$ .  $T^C$ , on the other hand, is the type space representing the agent's competence.

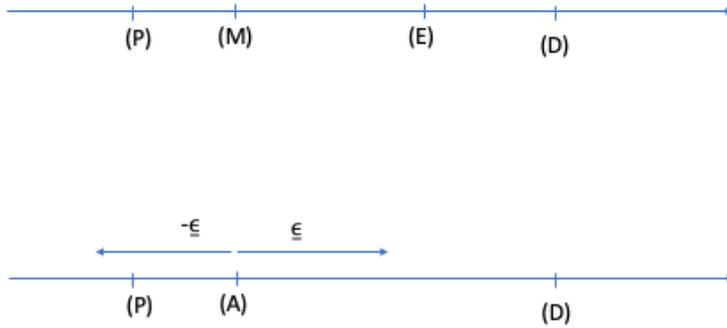
We assume that competent agents can effectively detect signals about the direction of random distortions to policy. For simplicity, in the baseline model, they observe the noise with which policy is implemented perfectly and can correct for it. Noise is represented by  $\omega \in \{-\epsilon, +\epsilon\}$ . With probability  $\frac{1}{2}$  the shock is  $-\epsilon$  and with probability  $\frac{1}{2}$  it is  $+\epsilon$ . A competent agent learns  $\omega$  with perfect precision, as in the standard delegation model. We assume that an agent is competent with probability  $\mu \in (0, 1)$  and incompetent with probability  $1 - \mu$ . We assume that  $0 < x_M < \epsilon < x_E < 1$ .

We take a moment to explain here why extremity and moderation are orthogonal to competence and loyalty. This may seem somewhat counterintuitive in light of the loyalty-competence trade-off discussed above. If it is indeed a “trade-off”, the agent hired by the former authoritarian regime for loyalty should also be lacking in competence and the agent hired by the former authoritarian regime for competence be lacking in loyalty? Why do we need two independent typespaces?

The reason lies in regime change. An agent hired by a former autocrat for competence, but not loyalty, may have an ideal point proximate to the new democratic Politician. But even an agent hired despite a lack of competence for his loyalty by the former autocrat may have acquired the necessary expertise while on the job and be potentially valuable to the new democracy.

To see this, suppose the preferences of the autocrat and the democratic politician succeeding him in office are as represented in Figure 2 below.

Let  $P$  represent the ideal point of whoever is setting the policy of law enforcement in the



**Figure 2:** *Reinterpretation of the loyalty-competency trade-off in post-authoritarian setting*

new democracy,  $D$  the ideal point of the former dictator, while  $M$  and  $E$  represent possible ideal points of the agents.  $E$  then is the agent ideologically proximate to the dictator but extreme from the point of view of the democratic politicians.  $M$  describes the ideology of the agent who is distant from the point of view of the dictator, but moderate from the point of view of the new democracy. Competence, represented in the lower panel of the figure, can be thought of as the ability of the agent to respond to the shock  $\epsilon$ , which can shift policy to the left or to the right. Note, that in an authoritarian setting, the loyalty competence trade-off could also be interpreted as follows: Sometimes an agent will be appointed at  $M$ , because he is competent. Yet that same agent, is at once loyal and competent from the point of view of the new democracy. Moreover, those are exactly the kinds of agent that politicians responsible for the reform of the police apparatus will want to retain.

Both typespaces are private information to the agent. The Politician can pay a cost and uncover the loyalty and competence (this is the interpretation of a selective purge). This cost is interpreted as taking the time to run a verification commission that studies and evaluates

the personnel files of secret police officers. These files contain relevant information about their performance, whether they received awards or sanctions, their tenure on the job as well as career trajectory, transfer history and education acquired before and during their service in the secret police.

For the baseline model, we analyze the following interaction between the verification commission,  $P$  and law enforcement agent,  $A$ .

In the first period, Nature determines if the agent is competent and whether his ideal point is moderate or extreme. In terms of Figure 2, the extreme agent has an ideal point close to but less than 1 than the moderate agent, while the moderate agent has an ideal point closer to 0 than the extreme agent.

In the second period,  $P$  (not having observed the distortion to policy that will result from the shock,  $\omega$ ) chooses one of three actions  $a_P \in \{tp, sp, np\}$ :

- thorough purge ( $tp$ ), fires everyone and appoints a replacement who shares an ideal point with  $P$ , i.e., 0, but has no competence (so cannot observe  $\omega$ );
- selective purge ( $sp$ ), where after paying cost  $v$  he can learn the competence and ideal point of the agent and only fire the extreme incompetent one;
- no purge ( $np$ ), where nobody is fired and no new information is learned.

In the third period, the original agent (if he was not replaced), implements policy. If the agent was replaced, then his replacement implements the policy. The game ends and payoffs are collected.

Utility functions are defined by distance of implemented policy (that is,  $p + \omega$ ) from players' ideal points. In the case of  $P$ , utility functions are also determined by the cost of conducting a selective purge (if one is chosen).

$$u_A(p; a_P) = -(|p + \omega - x_T|)$$

, where  $T \in \{M, E\}$

$$u_P(p; a_P) = -(|p + \omega|) - v * \mathbb{1}_{sp}$$

, where

$v$  represents the cost of appointing a verification commission that will conduct a careful analysis of personnel files before reaching each individual decision.

Since this is an extensive game of complete but imperfect information, we solve it for Subgame Perfect Equilibrium.

First, consider actions of the agent, starting with the replacement agent, which is appointed in the event of a thorough purge or if a selective purge reveals the agent is extreme (and so would be disloyal to the new democratic administration). This agent's ideal point is at 0, so her utility, as a function of her policy choice can be written as:

$$u_A(p; np) = -(|p + \epsilon|) * \frac{1}{2} - (|p - \epsilon|) * \frac{1}{2}$$

Given this agents' utility function, his optimal choice if  $p^* = 0$ .<sup>5</sup>

Consider now the agent without competence. Like the replacement agent, he cannot observe  $\omega$ , so he maximizes his utility function

$$u_A(p; np) = -(|p + \epsilon - x_T|) * \frac{1}{2} - (|p - \epsilon - x_T|) * \frac{1}{2}$$

leading him to choose  $p^* = x_M$  for the Moderate and  $p^* = x_E$  for the Extreme incompetent types, respectively.

Finally, consider the competent types of agents, who are aware of the value of  $\omega$ . Their maximization problem, in light of knowing  $\omega$  leads them to choose

---

<sup>5</sup> Technically, any  $p \in (-\epsilon, \epsilon)$  is optimal but changing the utility function to quadratic (see appendix) makes  $p^* = 0$  unique Gehlbach (2021).

$$p^* = \begin{cases} x_T + \epsilon & \text{if } \omega = -\epsilon; \\ x_T - \epsilon & \text{if } \omega = +\epsilon. \end{cases}$$

We can next move to  $P$ 's best response, whose expected utility of can be written as:

$$U_P(p^*; nocost, tp) = -\epsilon$$

This is a straightforward consequence of the fact that a thorough purge results in the appointment of the replacement agent who plays his optimal strategy described above as  $p^* = 0$ .

$$U_P(p^*, nocost, np) = \underbrace{-\theta\mu x_M}_{\text{moderate, competent}} + \underbrace{-(1-\theta)\mu|x_E|}_{\text{extreme, competent}} + \underbrace{\theta(1-\mu)(-|x_M + \epsilon| * \frac{1}{2} - |x_M - \epsilon| * \frac{1}{2})}_{\text{moderate incompetent}} + \underbrace{(1-\theta)(1-\mu)(-|x_E + \epsilon| * \frac{1}{2} - |x_E - \epsilon| * \frac{1}{2})}_{\text{extreme incompetent}}$$

This expression reflects what happens if no purge takes place. Starting from the top, the moderate and competent former secret police agent will adopt the policy that “absorbs the shock,”  $\omega$  (observed by the competent agent), to arrive at his ideal point,  $x_M$ . There are  $\theta * \mu$  (for moderate times competent) such cases. Moving on to the incompetent agents, reflected in the next two lines, recall that these agents too attempt to arrive at their ideal point but are unable to absorb the shock,  $\omega$  which with probability  $\frac{1}{2}$  makes them veer right from their ideal point and with the same probability makes them veer left. There are  $\theta(1 - \mu)$  such moderate agents and  $(1 - \theta)(1 - \mu)$  such extreme agents.

The above expression reduces to:

$$U_P(p^*, nocost, np) = -\epsilon(\theta(1 - \mu)) - \theta\mu x_M - x_E(1 - \theta) \quad (1)$$

Finally, the expected utility from the selective purge, so after paying a cost for a methodical verification process, is given by:

$$\begin{aligned}
 U_P(p^*, \text{paycost}, sp) = & \underbrace{-\theta\mu x_M}_{\text{moderate, competent}} \\
 & \underbrace{-(1-\mu)(\epsilon)}_{\text{replaced any incompetent}} \\
 & \underbrace{-(1-\theta)\mu\epsilon}_{\text{replaces extreme competent}} - v
 \end{aligned}$$

To understand this expression, recall that the selective purge reveals complete information on both the competence and the ideology of the former agents.

The above expression reduces to:

$$U_P(p^*, \text{paycost}, sp) = -\epsilon(1-\theta\mu) - \theta\mu x_M - v \tag{2}$$

What does  $P$  choose as her optimal action? The optimal policy can be easily gleaned from comparing the payoff from a thorough purge,  $-\epsilon$ , to (1) and next, (2).

Note first, that, just as the payoff from a thorough purge, the entirety of (1) is negative. However, the coefficient on  $\epsilon$  is less than 1, so at least for some parameter values, both (1) and (2) can be greater than  $-\epsilon$ .

For instance, for both (1) and (2), an ideal point of the moderate type of agent (the one with an ideal point closer to the Politician) unambiguously increases the value of expression (1), which is consistent with standard delegation framework: If the agent's and principal's preferences are aligned relative to the uncertainty associated with policy implementation, then delegation is more likely. What makes (1) more likely to be smaller relative to (2) and  $-\epsilon$ , in turn, is  $x_E$ , the position of the extreme agent.

Another factor that will decrease (1) relative to  $-\epsilon$  is the proportion of extreme agents  $(1-\theta)$ . The effect of  $\mu$ , the proportion of competent agents, in (2), is intuitive: A higher proportion of competent agents increases (2). This can be seen by comparing the first two

expressions in (2):  $-\epsilon(1 - \theta\mu)$  to  $-x_M\theta\mu$ . Since  $\epsilon > x_M$ , the positive effect of  $\mu$  on (2) outweighs the negative effect.

To arrive at our proposition, we proceed with a more systematic analysis. Recall, that ultimately, we are interested in uncovering when a selective purge is more attractive than both a thorough purge and a “no purge.” This question will produce the comparative statics we will eventually be testing with our paired comparison approach. The payoff associated with a thorough purge is flat at  $-\epsilon$ , as it does not depend at all on  $\mu$  or  $\theta$ . Conditions for when when a selective purge is preferred over a thorough or “no purge” can be extracted from (2) and (1) as:

$$\begin{cases} -\epsilon < -\epsilon(1 - \theta\mu) - \theta\mu x_M - v \\ -\epsilon(\theta(1 - \mu)) - \theta\mu x_M - x_E(1 - \theta) < -\epsilon(1 - \theta\mu) - \theta\mu x_M - v \end{cases}$$

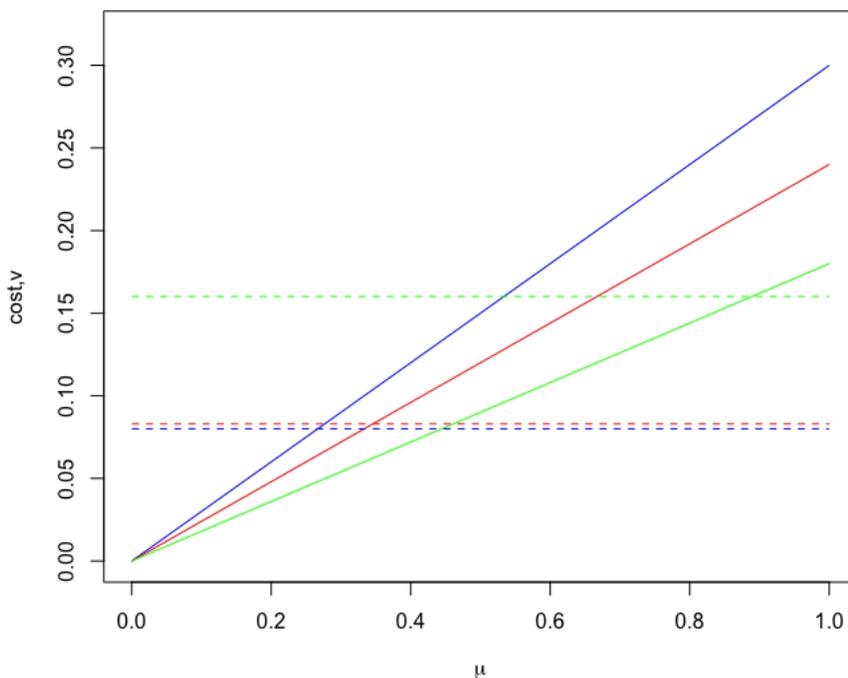
After simplifying the above expression we arrive at our proposition (see Appendix for all calculations.)

**Proposition 1.** *The Politician chooses a selective purge over “no purge” or thorough purge if and only if the cost of verification,  $v$  is sufficiently small, that is,  $v < v^*$ , where  $v^* = \min\{(x_E - \epsilon)(1 - \theta), \mu\theta(\epsilon - x_M)\}$*

The threshold that has to be cleared by  $v$  sometimes, but not always, increases with  $\mu$ , the proportion of competent agents. To see this, note that as long as both conditions described in (1), are for some parameter values, binding constraints, the choice of a selective purge may no longer be responsive to  $\mu$  in the same way it is responsive to  $\theta$ , the proportion of extreme agents in the pool.

A further comparison of the two constraints shows that for  $\mu > \mu^* = \frac{(x_E - \epsilon)(1 - \theta)}{\theta(\epsilon - x_M)}$ , the constraint  $v < (x_E - \epsilon)(1 - \theta)$ , which corresponds to “no purge” is binding, while for  $\mu < \mu^* = \frac{(x_E - \epsilon)(1 - \theta)}{\theta(\epsilon - x_M)}$ , the constraint that is binding corresponds to “thorough purge” and is expressed by  $v < \mu\theta(\epsilon - x_M)$ .

This is intuitive: when  $\epsilon$  is low, the value of competence is negligible, making a thorough purge a potentially attractive option. Not much is lost from replacing experienced agents of enforcement with new employees if the cost of making a mistake is small. When  $\epsilon$  is high, on the other hand, “no purge” is relatively more attractive because lack of competence hurts the new democracy more than agents with misaligned preferences. This is further illustrated in Figure 3.



**Figure 3:** Constraints on  $v$  for selective purge to be optimal as a function of  $\mu$  for two sets of parameter values (Blue:  $\epsilon = .7, x_M = .2, x_E = .9, \theta = .6$ ; Red:  $\epsilon = .5, x_M = .2, x_E = .9, \theta = .8$ ); Green:  $\epsilon = .5, x_M = .2, x_E = .9, \theta = .6$   
Note: dotted blue and red lines overlap completely so have been artificially separated by a hair to make them both visible.

The solid lines in the figure correspond to the constraint from thorough purge and the dotted lines correspond to the constraint from no purge.

Beginning with constraints represented in blue ( $\epsilon = .7, x_M = .2, x_E = .9, \theta = .6$ ) representing relatively high demand for expertise (as the parameter responsible for policy shocks,

$\epsilon = .7$ ), we see that for relatively low values of  $\mu$ , the binding constraint comes from the thorough purge option: the solid line is below the dashed line. This binding constraint holds until about .24, at which point the binding constraint becomes the one corresponding to the no purge option (the dashed line in blue). Once  $\mu$  assumes the value of .24 for this set of parameter values, the cost of verification must be below .08, regardless of any subsequent change in  $\mu$ . The point where the solid line crosses the dashed line determines which expression defines the binding constraint. The green line represents more predictable (with  $\epsilon = .5$ ) conditions relative to the previously discussed blue, while all other parameters remain the same. With more certainty, the constraint representing the thorough purge is the main constraint for longer: the solid line does not intersect the dashed line until  $\mu$  reaches about .85. Finally, the red constraints correspond to the same uncertainty as green (with  $\epsilon = .5$ ), but with a higher probability of the agent being moderate ( $\theta = .8$ ). The higher  $\theta$  decreases the slope in the line corresponding to the “thorough purge” constraint, so that it remains binding for a higher proportion of  $\mu$  (competent agents) than the blue parameters but a lower proportion than the green parameters.

The parameters for the three illustrative cases presented here have been chosen to isolate the increase of  $\epsilon$  (the change from the green to the blue line) and the change in  $\theta$  (the change from green to red). It is worth pointing out that the increase in  $\epsilon$  (the demand for competence) makes the selective purge at least just as valuable and for low competence levels more valuable.

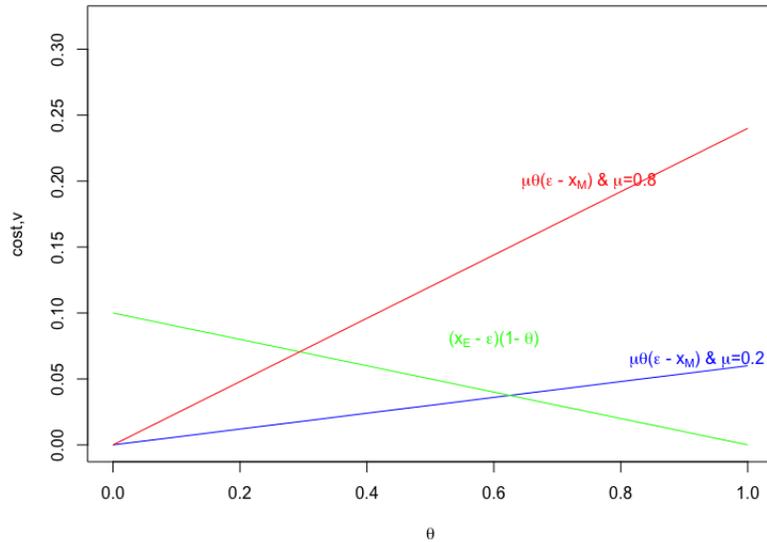
Meanwhile, the change in the proportion of moderate agents from .6 (green) to .8 (red) makes the condition for conducting a selective purge easier to satisfy for low levels of competence, but harder to satisfy for a high level of competence.

To see this last point in another way, consider Figure 4, which illustrates the same constraints, though as a function of  $\theta$  for two different values of  $\mu$ , .2 (blue) and .8 (red). These values are only relevant for the constraint  $v \leq \mu\theta(\epsilon - x_M)$ . The green line represents the constraint that corresponds to  $v \leq (x_E - \epsilon)(1 - \theta)$  and does not contain  $\mu$  at all, while

$x_E = .6$  and  $\epsilon = .5$  (the same as in the red and blue lines).

The point of this figure is to show that for a low proportion of competent agents, the binding constraint on  $v$  that has to be cleared for a selective purge to be attractive is increasing in  $\theta$ , the proportion of moderate agents, while for a high proportion of competent agents, the binding constraint is decreasing in  $\theta$ . This can be seen from the fact that the blue line remains below the green line until, roughly  $\theta = .65$ . After the blue and green lines intersect, the binding constraint becomes the one represented by the green line and it declines with  $\theta$ , that is, a higher proportion of moderate agents makes it harder to satisfy. This is intuitive: When most agents are moderate, the risk from no purge is low enough that the cost of a selective purge must be very low to make it worthwhile (note, that decreasing  $\epsilon$  or increasing  $x_M$  could shift that point of intersection further to the right).

Consider now the case where  $\mu = .8$ , represented by the red line. High  $\mu$  makes this line steeper and intersects the green line for a smaller  $\theta$ , approximately at  $\theta = .35$ . Hence for high proportions of competent agents, the effect of  $\theta$  is to mostly make selective purges less attractive. Only for very  $\theta$  (below  $.35$ ) does  $\theta$  have a positive effect on selective purges.



**Figure 4:** Constraints as function of moderate agents

The implications from the above analysis are summarized in Table 1 and are the basis of our empirical testing strategy executed in the following section.

interpretation	low competence	high competence
range of $\mu$	$\mu < \frac{(x_E - \epsilon)(1 - \theta)}{\theta(\epsilon - x_M)}$	$\mu \geq \frac{(x_E - \epsilon)(1 - \theta)}{\theta(\epsilon - x_M)}$
Binding constraint on $v$	$v < \mu\theta(\epsilon - x_M)$	$v < (x_E - \epsilon)(1 - \theta)$
Optimal ordering of verification commission	$sp \succcurlyeq tp \succcurlyeq np$	$sp \succcurlyeq np \succcurlyeq tp$

**Table 1:** *When are selective purges preferred?*

In a nutshell, the predictions of the baseline model diverge depending on whether the proportion of competent agents is high or low. When the proportion is low ( $\mu < \frac{(x_E - \epsilon)(1 - \theta)}{\theta(\epsilon - x_M)}$ ), an increase in  $\theta$  makes the condition for selective purges to be chosen easier to satisfy. Hence, we expect that in regions with low average competence, a higher proportion of moderate agents will make selective purges more likely. Note that the last row of Table 1, shows which kind of purge should be selected if the selective purge is not available. Since the binding constraint is the one derived from comparing the payoffs to a thorough purge with the payoffs of a selective purge, the ordering is  $sp \succcurlyeq tp \succcurlyeq np$ .

In contrast, when the proportion of competent agents is high, the binding constraint on  $v$ , the cost of a selective purge is a decreasing function of  $\theta$  ( $v < (x_E - \epsilon)(1 - \theta)$ ). Hence an increase in the proportion of moderate agents makes the condition harder to satisfy. This means that in regions with high proportions of competent agents, an increase in extreme agents will make selective purges more likely.

The last row of Table 1 indicates that because the binding constraint is derived from comparing the payoffs to selective purge with payoffs to no purge, the optimal ordering of preferred purges is  $sp \succcurlyeq np \succcurlyeq tp$ .

In the following section, we take the main findings of the model to data. To summarize, the most robust finding is that the only parameter of the model that unconditionally makes a selective purge more attractive is a decrease in the costs of setting up a verification commission.

We also find that the competency of former agents moderates the effect of ideological extremism on choosing selective purges. When competence is low, more extreme agents make selective purges less likely. But when competence is high, more extreme agents make selective purges more likely.

This second pair of implications will guide our construction of the empirical test based on paired comparisons (Tarrow 2010). First, however, in the next section, we explain why data on the operation of verification commissions in 1990 from the Polish IPN is ideal for our purposes.

## 4 Verification of security officers in Poland

Several reasons make Poland an appealing case to study the turnover of agents of repression accompanying regime change.

First, during its transition years, Poland had to deal with security forces whose loyalties could potentially destabilize the state. In the 1990's private protection and security services were highly sought after. Urban areas became hotbeds of violence committed by actors ranging from mobs of organized crime to soccer fans. In this context, the new democratic government had to decide how to retain competent moderate officers capable of staving off the threats of the transition while sifting out the disloyal and/or possibly incompetent "bad apples."

Poland's was not just a democratic transition but also a transformation of a socialist-planned to a free-market economy. The economic transformation was abrupt. Dubbed the "big-bang," it created ample opportunities for those with access to information for fast enrichment. The political transition was slow and protracted. Over two months of "Roundtable Talks," preceded by secret preparations negotiated in a Warsaw suburb, the old government negotiated the terms of including the Solidarity trade union in the formal political decision-making.<sup>6</sup> Ultimately, Solidarity candidates were allowed to run for 35% of seats to the Sejm

---

<sup>6</sup> Despite its name, the Roundtable negotiation was bilateral.

FIGURE 5: *Example of a Regional Verification Commissions summarizing the commission's decisions*

SKŁAD KOMISJI KWALIFIKACYJNEJ DLA FUNKCJONARIUSZY  
BYŁEJ SŁUŻBY BEZPIECZEŃSTWA

województwo ..... GDAŃSK .....

przewodniczący	.....	Franciszek JAMROŹ.....
przedstawiciel Szefa Urzędu Ochrony Państwa	.....	Adam HODYSZ.....
przedstawiciel Komendanta Głównego Policji	.....	plk Stefan BACHORSKI.....
przedstawiciel Zw.Zaw.	.....	por. Jerzy PELC.....
poseł	.....	Jan Krzysztof BIELECKI.....
poseł	.....	Czesław NOWAK.....
senator	.....	Lech KACZYŃSKI.....
osoba o uznanym autorytecie	.....	.....
sekretarz	.....	.....

- Komisja rozpoczęła działalność w dniu .....

- Zakończono działalność w dniu ..... 25.07.1990r .....

- Kwalifikacja objęto .. 439 .. osób

- Zaopiniowano: pozytywnie ..... 405 ..... osób  
negatywnie ..... 34 ..... osób  
odwołało się ..... 34 ..... osób

rozpatrzono pozytywnie ..... 11 ..... odwołań  
negatywnie ..... 20 ..... odwołań

time deliberating *and* verified a high proportion of officers negatively; and *no purge* as an instance where the RVCs took little time deliberating *and* cleared a high proportion of officers to serve under the new democratic system. Thus, no purge means that the commission was relatively lenient, not necessarily that all the agents were verified positively. Similarly, a thorough purge means that the commission was relatively stringent.

Demarcating the categories of “less” and “more” time into clear categories requires a definition of cutoffs. In Table 2) we operationalize it as above (more) and below (less) the mean number of hours expended per file.

This rule allows us to classify the vast majority of wojewodztwa according to the type of purge they used in dealing with former secret police agents (Figure 7, Figure 6)<sup>9</sup>.

Our classification necessarily rests on arbitrary choices concerning the levels of key vari-

<sup>9</sup> The remaining wojewodztwa could not be classified as the dates of operation of the RVCs (and hence the time devoted to each application) are missing from the files.

FIGURE 6: *Classifying purge types*

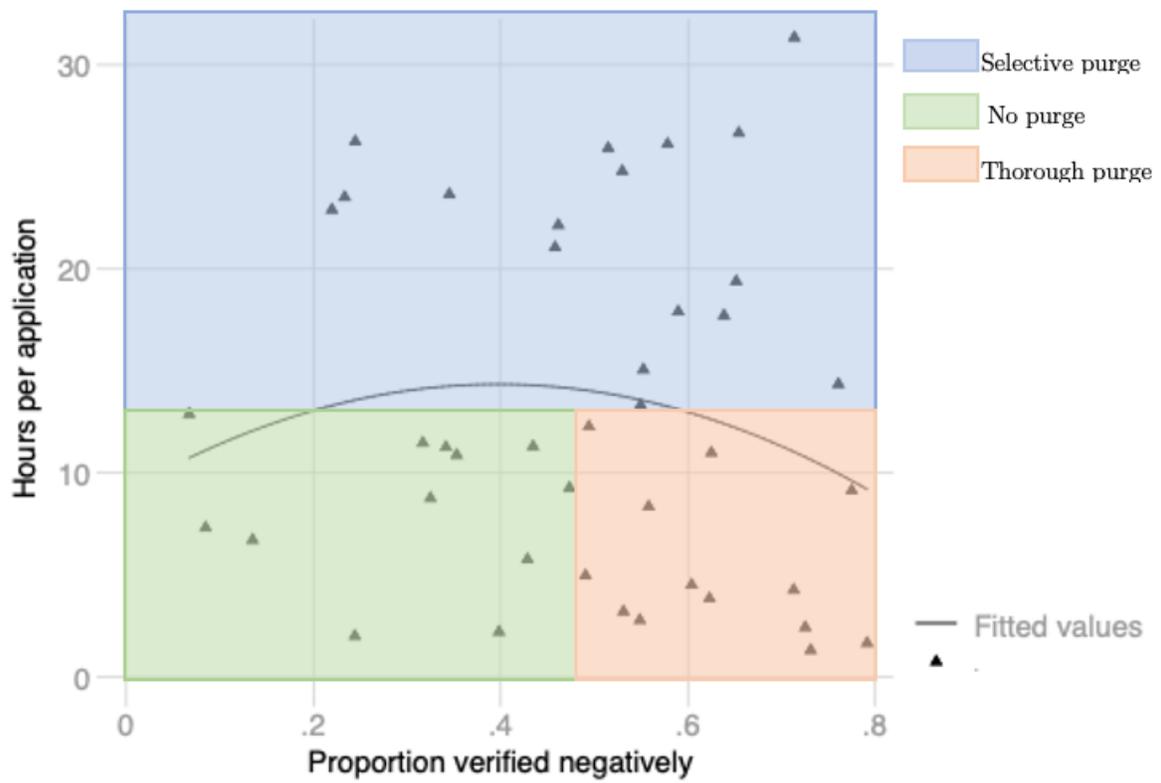


TABLE 2: *Types of purges*

Purge type	Proportion verified negatively	Hours per application	Frequency	Percent
No	<0.48	<12.9	12	28.57
Selective		>12.9	17	40.48
Thorough	>0.48	<12.9	13	30.95

ables. To increase its credibility, we corroborate it with supporting information from the archives of RVCs, consisting of letters and reports that allow us to paint a richer picture of the commissions' work.

In wojewodztwa classified as having no purges, the documents consist of complaints regarding the low levels of purges. For instance, Wloclawek's Solidarity Trade Union rep wrote a letter to the RVC complaining about the high level of positive verification in the wojewodztwo (BU/3546/52). In Zamosc, the commission wrote "Society judged our position as too lenient. SB's activity in the Zamosc area was particularly brutal and repressive." (BU/3546/53 p.22). Finally, in Radom, the leader of the RVC wrote "We feel that the Commission's activities have not met public expectations." (BU/3546/39). Documents from wojewodztwa that saw selective purges show fewer complaints. In Bydgoszcz, the commission reports "considerable public approval of the work of the RVC. Attitudes demanding revenge soon morphed into acceptance and even empathy." (By/453/47 p.204) In Tarnow, the selective aspect is also clearly visible: "(i)n cases where the in-depth information did not give unequivocally negative indications, the RVC requested positive consideration of the appeals". At the same time, Tarnow's RVC was fully convinced that "the persons assessed negatively and whose decision was upheld should not work in the ministry (...)." (BU/3546/48 p.8).

Correspondence stored in the archives of the wojewodztwa that underwent thorough purges strikes a yet different tone, suggesting a desire for revenge among commissioners. According to the author (a former SB officer) of a letter addressed to the Ombudsman for Human Rights "on its first day of proceedings, the commission [RVC in Tarnobrzeg] handled

a vast majority of valuable and highly educated personnel. (...) Moreover (...) there was a list of 40 names drawn up by “Solidarity” demanding they be dismissed. Among those negatively verified were secretaries, an archivist, and even an officer who served on one of the new parliamentary committees (...) This was but another act of revenge carried out in our country.” (BU/3546/47 p.41).

As a final robustness check of our categorization, we use the proportion of officers that appealed their RVC decision along with the percentage of those who were granted appeal. If our purge decisions have been correctly classified, the proportion of appeals should be highest in those wojewodztwa that we classified as implementing thorough purges while wojewodztwa with no purges should have the lowest proportion of appeals. Finally, the thorough purge wojewodztwa, should have most reversals of negative decisions.

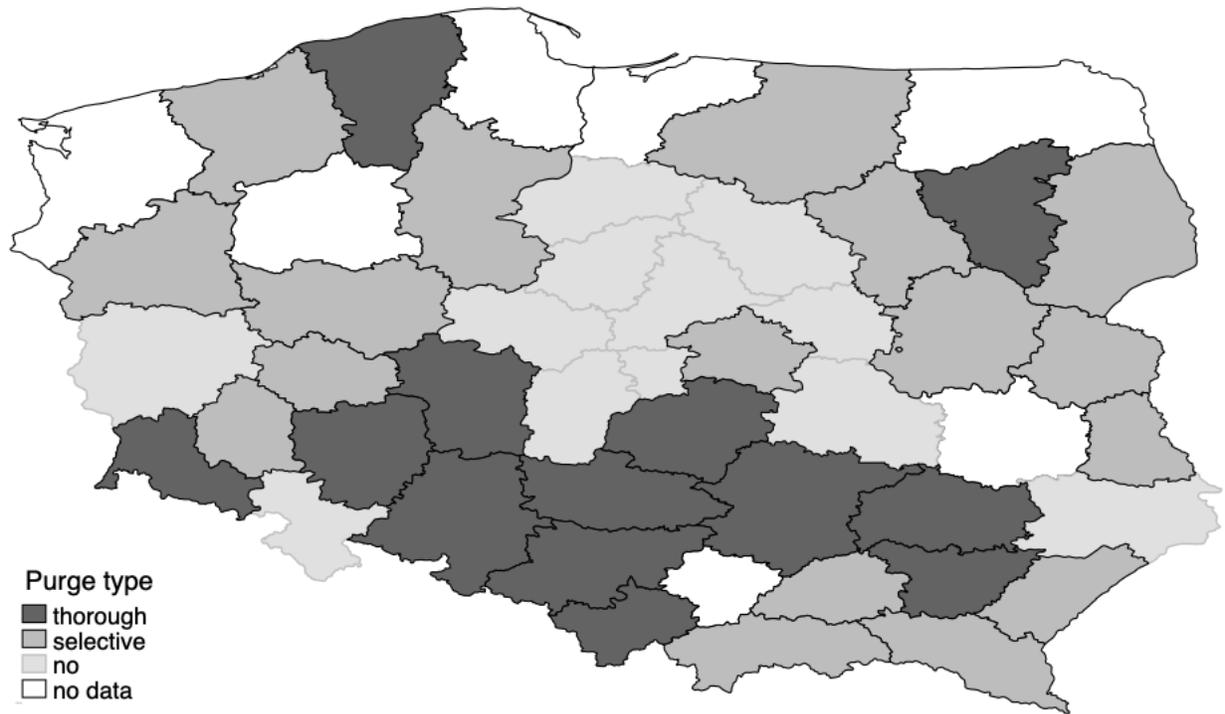
A comparison of means shows exactly that: In wojewodztwa with thorough purges 91.9% officers verified negatively had appealed the result, while in wojewodztwa with selective purges the corresponding number was 90.4%. Finally, wojewodztwa with no purges had only 85.4% of the rejected officers appeal their decisions. Of these appeals, the highest proportion— 75.5%— of reversals took place in wojewodztwa experiencing thorough purges, followed by 60.7% in wojewodztwa with selective purges. The corresponding figure for wojewodztwa with no purges was 58.7% .

These additional robustness checks increase the credibility of our classification rule. Moreover, the results of our analysis as well as the appeals data reported above are robust to using median, rather than mean, values of the proportion of negatively verified and time spent per application. As expected, using the median classifies relatively more wojewodztwa as experiencing a selective purge (see Figure 7).

Having discussed the dependent variable, the key explanatory variables for testing our hypotheses are:

- the level of competence among officers in a given wojewodztwo;
- the extent of extremism among the officers;

FIGURE 7: *Classification of purge types in wojewodztwa*



- the cost of conducting a selective—that is, associated with a careful reading of personnel files—verification commission

We discuss each in turn. First, to operationalize competence of the security services officer corps we draw on data from an IPN publication covering the basic statistics on the type and number of officers at the regional level (Piotrowski 2003). The most relevant information is on the proportion of officers tasked with securing economic resources and employed in the Department for the Protection of the Economy (PoE, *Wydział Ochrony Gospodarki*) in 1989. We use these figures as a proxy for competence, as the prestige of this department attracted highly capable officers (Kozłowski 2019), with skills that can be no doubt considered “usable” Grzymala-Busse (2002) by the new democratic regime. In contrast to officers who spent their time repressing the anticommunist opposition or the Catholic church, officers from PoE were trained to trace crimes that were more economic than political. According to our data, the proportion of PoE officers among all officers employed in the region varied between 0.13 and

0.53 (see Table 3).

Second, to capture the level of extremism among officers we use the intensity of repression during the Polish Martial Law (1981-1983) broken down by wojewodztwo. The level of repression is tricky to measure because of problems of reverse causality: Repression is a response to anti-governmental action which when effective stifles the demand for anti-governmental action, and—consequently—repression Ritter and Conrad (2016). To avoid such issues, we begin with calculating the number of people convicted of political crimes following the implementation of Martial Law, but then divide it by the number of people that were brought to court (Instytut Pamięci Narodowej 2021) in each wojewodztwo. Standardizing the number of convictions by the number of those tried allows us to establish the severity of political repression in each wojewodztwo controlling for the level of dissent.

We believe this is an adequate measure of the proportion of extreme agents, as it captures the response of a local security apparatus holding the level of opposition constant. On average, 63.9% of those tried were sentenced for political crimes. However, this value ranges from 0% to 100% across wojewodztwa, suggesting considerable variation in how active the judiciary was in exercising its ability to punish dissidents. While the courts and the security services were nominally separate organizations, research by Popova (2012) and McCarthy (2015) suggests that in autocracies, the actions of these institutions are more rigidly aligned. In light of these considerations, higher levels of repression by courts can be used as a proxy for extremism among the local security services.

A possible limitation of this measure is that it is agnostic with respect to the number of members of the opposition. Hence, as a robustness check, in the analysis that follows we also consider an alternative measure of repression based on the ratio of those sentenced as a proportion of Solidarity members in a given wojewodztwo.

Finally, to account for the cost of running verification commissions we rely on population density in 1985.<sup>10</sup> We argue that the costs of selective purges were related to demands

---

<sup>10</sup> 1985 is simply the midpoint between two focal moments of the communist regime in Poland: the implementation of Martial Law in December 1981 on the one hand and the first partially free elections

TABLE 3: *Descriptive statistics*

Variable	Obs	Mean	Std. Dev.	Min	Max
Verified negatively	49	0.477	0.198	0.067	0.791
Hours/application	42	12.90	8.64	1.32	31.35
Population	49	761814	596134.1	244300	3916400
Population density	49	140.74	145.91	42.80	748.75
Extremism	47	0.639	0.26	0	1
Competence	48	0.27	0.09	0.13	0.53

for justice (or revenge) from citizens who had been directly wronged by the secret police. Intuitively, more urban wojewodztwa with fewer rural areas had a denser network of dissident activists. Consequently, the probability in more urban wojewodztwa, that any given resident personally knew someone persecuted by secret police officers is higher than in rural wojewodztwa.

Building on these insights we assume that in wojewodztwa with a smaller population density the commission would pay a lower cost for engaging in a scrupulous examination of individual officer files. In wojewodztwa with a higher population density, the easy way out was to succumb to popular demand for revenge.

An alternative justification for using population density as a proxy for the price of a selective purge is to note that the composition of RVCs was not in any way tied to the number of officers employed by SB or to how many of them applied to be verified. At the same time, more densely populated areas would have higher numbers of SB officers. Consequently, the time any given RVC member could spend on a file decreased with population density, increasing the cost of a selective purge also.

Finally, we interpret the policy shock  $\epsilon$  as the uncertainty related to the regime transition and hence common for the entire country. This implies that in our theoretical proposition,  $\epsilon$  can be held constant and so does not enter directly into our empirical analysis and does not need to be operationalized.

---

in June 1989.

A summary of how we operationalize the key parameters of our model is as follows:

- $\mu$  - the proportion of competent agents, is measured as the proportion of officers who before the transition worked in the Department for the Protection of the Economy
- $v$  - the cost of setting up the commission is proxied for by the population density in each wojewodztwo
- $\theta$  - probability that the agent is moderate. Conversely,  $(1-\theta)$  is the probability that the agent is extreme, which we measure using the level of repression in each wojewodztwo

## 4.2 Analysis

Having established the empirical interpretation of the model's key parameters, we are ready to present our empirical hypotheses. Based on the empirical implications derived from our model in section 3 we should see:

1. More selective purges where the cost of setting up a verification commission, measured by population density is small.
2. In wojewodztwa where many secret police officers were employed in PoE (*high competence*) an increase in *extremism* of a typical officer increases the chances of a selective as opposed to no purge.
3. In wojewodztwa where few secret police officers were employed in PoE (*low competence*) an increase in *moderation* of a typical officer increases the chances of a selective as opposed to thorough purge.

Our empirical tests are a combination of medium-N analysis using comparisons of means and a paired comparisons approach (Tarrow 2010). Instead of bemoaning our low number of observations ( $N=42$ <sup>11</sup>), we see this as an opportunity to introduce yet another research design

---

<sup>11</sup> While we have data on the proportion of negatively verified officers for all wojewodztwa, not all the files included the period of operation of the verification commission

to the toolkit of scholars doing work in the spirit of Empirical Implications of Theoretical Models.

First, via a comparison of means in Table 4 we address the model’s predictions concerning the effect of the cost of setting up a verification commission on the choice of a selective purge.

TABLE 4: *Means of the purge type determinants*

Type of purge	$\mu$	SD( $\mu$ )	$(1 - \theta)$	SD( $1 - \theta$ )	$v$	SD( $v$ )
No	0.280	0.087	0.709	0.277	196.63	234.82
Selective	0.257	0.066	0.544	0.250	86.45	34.72
Thorough	0.239	0.109	0.687	0.246	160.82	136.61

As postulated by the model, wojewodztwa that chose a selective purge are also the ones where the verification cost ( $v$ ) was relatively low: those that have seen selective purges are characterized by lower average population density, that is a lower cost of conducting a thorough investigation was less costly.

### 4.3 Paired comparisons

The last two theoretical predictions of our model concern instances where the cost or a selective purge exceeds the threshold to make it worthwhile for the RVC. In those circumstances, the best choice will be determined by the interaction of competence and ideological moderation of agents.

When both competence and moderation are low, a selective purge is too expensive and a thorough purge is chosen. On the other hand, when both competence and moderation are high, the pool of agents is of high quality and, rather than paying for a selective purge, the verification commission can choose no purge. Simply put, the proportion of extreme agents does not have a linear effect on the likelihood of a selective purge, but is instead conditional on competence: in high competence settings selective purges were more likely where agents were more extreme, but where competence was low, selective purges were likely where agents were more moderate.

TABLE 5: *Wojewodztwa compared and their characteristics*

Wojewodztwo	Repression	Population density	Competence	Purge type	Officers verified
Tarnowskie	0.39	154.5	0.211	selective	139
Rzeszowskie	0.71	157.2	0.213	thorough	206
Bydgoskie	0.54	104.7	0.363	selective	284
Wloclawskie	0.44	96.8	0.354	no	164

These key mechanisms cannot be easily studied in a large-n framework. First, the conditional effects would require us to split the already limited sample; Second, and more importantly, however, our key interest lies in confirming that the mechanisms that underpin the quantitative data are what we posit them to be. Hence, we zoom in on two pairs of wojewodztwa illustrating the mechanisms discussed in our theory.

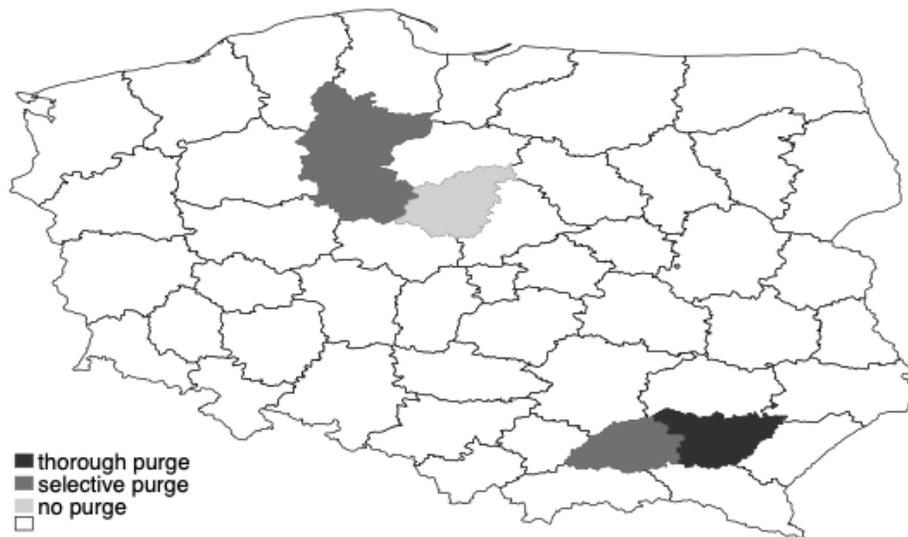
The paired comparison approach, which we used to select two pairs of wojewodztwa, is central to our testing strategy. The method is an adaptation to the social sciences of John Stuart Mill (1869)'s Method of Difference (Most Similar System Design). It enables linking the key explanatory variable to the dependent variable by choosing cases that are most similar on the relevant control variables. Hence, in spirit, it is akin to statistical matching.

To the best of our knowledge, this approach has never yet been used to test a formal model. Yet the suitability of the paired comparison approach to testing implications based on comparative statics of a formal model is striking. First, it allows us to control for variables operationalizing the model's parameters by using those variables to select cases. Second, many critics of EITM point out that even though a formal model isolates the effects of parameters *omitted* from the formal model, one cannot do the same when carrying the model's parameters over to a regression framework. There, any omitted variables can bias the results. By selecting our pairs of cases from the same country and region, we are effectively controlling for many hard to account for cultural or geographical determinants

and thus escaping the criticism that Sydney Tarrow (2010) levied against Skocpol (1979). Moreover, when using the conditioning parameter—the proportion of competent agents (or those employed in PoE)—we chose neighboring wojewodztwa to ensure that they shared the same imperial legacies.<sup>12</sup> In doing so, we were able to effectively control for unaccounted institutional legacies of the past, thus escaping the criticism Tarrow (2010) levied against the work of Putnam et al. (1994).

The wojewodztwa presented in Table 5 and Figure 8 were chosen to satisfy a Most Similar Systems Design. Specifically, the two pairs had similar levels of competence (agents in PoE), cost (population density) and shared historical and cultural legacies, as well as a host of other variables (Table 8). At the same time, in each pair, one wojewodztwo saw a selective purge while the other saw a thorough or no purge, respectively. This design allowed us to “control for” common systemic characteristics while inter-systemic differences play the role of explanatory variables (Przeworski and Teune 1981).

FIGURE 8: *Location of the compared wojewodztwa*



Consider first, in the top part of Table 5, the two wojewodztwa with below-average levels of PoE agents (competence), but similar population density: tarnowskie and rzeszowskie, two

<sup>12</sup> Poland’s territory for 123 years (the period of the so-called Partitions) was partitioned by three empires: Russia, Prussia and Hapsburg.

neighboring wojewodztwos in southern Poland, depicted also in Figure 8. These wojewodztwa saw different types of purges even though they are very similar in terms of population density: Their share of city and village gminas<sup>13</sup> is nearly identical (and above average) and they are similarly distant from Warsaw, leading to a similarly peripheral character. Moreover, they both lie in a part of Poland that was controlled by the Hapsburg Empire from the 18th century, potentially leading to comparable historical legacies in their institutions.

Yet, tarnowskie had a selective purge, where in 19 days the commission negatively considered the application of 39% of the 139 officers who sought verification. Rzeszowskie, on the other hand, had a thorough purge, with all 206 applications considered in 3 days and 54% verified negatively. The main difference between the two was the level of repression. Rzeszowskie experienced brutal repression throughout the 1980s. In the early days of the Martial Law around 100 activists were arrested, 163 more threatened. In the following years, security forces violently broke up anti-communist demonstrations (Gliwa n.d.). The difference in repression is clear also according to our quantitative measure that we use to operationalize extremism among the agents. In contrast to rzeszkowskie, in tarnowskie, the number of activists arrested and placed in internment camps was relatively low (Redakcja 2006). Hence, the expected proportion of extremist agents in rzeszowskie was higher than in tarnowskie, making tarnowskie a good candidate for a selective purge.

Our second pair of wojewodztwa are bydgoskie and wloclawskie. They share many important characteristics. First, according to the variables operationalizing parameters of our model, both wojewodztwa have a relatively high, above-median population density, making setting up a selective purge relatively costly. Second, they both had a similar (and high) proportion of agents employed in PoE (36% and 35%, respectively), which we use to operationalize competence. Beyond the variables operationalizing our model's parameters, bydgoskie and wloclawskie shared institutional legacies: as geographic neighbors, they remained under Prussian control for a significant part of the Partition period<sup>14</sup>. Bydgoskie

---

<sup>13</sup> Gmina is a smaller administrative unit, akin to a district.

<sup>14</sup> Although after the Congress of Vienna, Wloclawek, a small part of wloclawskie became part of Con-

and wloclawskie even have comparable levels of “Solidarity” membership: 27% and 21% in bydgoskie and wloclawskie respectively<sup>15</sup>. Where the two wojewodztwa differed, is in the levels of repression (measured by the proportion of sentenced dissidents)—our independent variable—and the type of purges they experienced—our dependent variable.

The RVC in bydgoskie chose a selective purge. As described above, in bydgoskie commission members devoted a considerable amount of time to careful verification, which suppressed discussions of retributive justice towards the SB (By/453/47 p.204). In contrast, in wloclawskie, the commission chose no purge, allowing almost all the candidates for verification through.<sup>16</sup>

We attribute the difference in purge types between the two wojewodztwa to the difference in agents’ extremism, as posited by our model. The higher level of extreme agents in bydgoskie than wloclawskie, while holding constant the high level of competence, according to our theory, made a selective purge comparatively more likely in bydgoskie where the cost of selective verification was worthwhile.

At the same time, our model allows us to account for an outcome that many commentators found puzzling: Bydgoszcz had been a site of one of the most brutal repressions during anti-communist resistance (before the introduction of Martial Law in 1981). The repressions culminated with severe beatings of “Solidarity” members that resulted in several hospitalizations (Chincinski 2002). The event led to unrest throughout Poland but particularly in Bydgoszcz. Had the RVC given in completely to calls for revenge, it would have no doubt picked a thorough purge. Instead, it chose a selective purge. Our model helps understand why retribution that seems to be a natural response to repression gave way to practicality.

---

gressional Poland in the Russian Partition

<sup>15</sup> While the national average was 23.9%, it ranged from 8.5% to 44.4%

<sup>16</sup> This caused an uproar when several Solidarity affiliates from around the country complained to the Ministry of Interior that the RVC had ignored their complaints that positive verification of any but two security agents was morally unacceptable in light of the persecution members of Solidarity faced from the hands of the SB (BU/3546/52 p.63-69). This quote suggests that Solidarity affiliates believed the RVC was ignoring public opinion, which is consistent with mechanisms we label as “no purge”.

## 5 Conclusion

Protection of the rights of life, liberty, property, and contract is a fundamental function of the state, even according to the most minimalist conception (Nozick 1978; Mack 2018). Governments must ensure that the institutions offering protection, such as the police, security services, and courts function properly. But how do new governments recovering from periods of authoritarian rule reform, repurpose or replace their security apparatuses without surrendering these functions of the state?

In this paper, we formalize the dilemma facing policymakers in the aftermath of the transition to democracy, when new democratic governments must rebuild their “ships while at sea,” to use the words of Jon Elster et al. (1998). Specifically, we analyze the policies of reforming and/or purging the security apparatus of the authoritarian regime.

Without rigorous analysis, the answer to the question of who should be fired seems obvious: the bad apples. But the meaning of “bad” here is ambiguous. It can refer to a lack of competence or it can denote preferences ideologically distant from the new democratic government.

Using a simple formal model, that placed familiar elements, such as the loyalty-competence dilemma, in the context of regime change, we found unique circumstances when new democracies ought to take the time to carefully verify their security apparatus through a process we have called a “selective purge.” We show that selective purges are more likely when the costs of verification are low. We also established an interactive effect between the extremism of former agents (or their loyalty to the previous regime) and competence. Specifically, in cases where the competence of the former agents was low, a decrease in extremism makes it more likely that a selective purge will be chosen over a thorough purge. However, when competence is high, an increase in extremism makes a selective purge more likely to be chosen over no purge.

Our theoretical findings are corroborated by a comparison of means analysis using data from the operation of 49 regional verification commissions in Poland in 1990 and with a

paired comparison approach. With this dual research design, we find corroborative support for our theory.

This simple model lends itself to several modifications that we plan to examine in the future. A plausible extension could give the agents a chance to appeal (that is accept or reject the decision to be purged) and have the appeal reconsidered by a “central” verification commission.

We conclude this version with consideration of scope conditions.

It is worth considering the extent to which Poland’s verification commissions are representative of screening mechanisms for former secret servicemen in the Post-Communist context. Poland shares with countries in the region the dual nature of regime change, encompassing both a transition from authoritarianism and a transition from a planned economy. What sets Poland apart from many countries in the region, however, is the very gradual process of the transition and the fact the pace of transition itself was the subject of negotiation.<sup>17</sup> The pace likely factored into the price of selective purges, making careful verification cheaper. One could also make the opposite argument according to which gradual transitions gave outgoing autocrats a chance to destroy incriminating personnel files of secret police officers, which would ultimately make the work of verification commissions harder. Kozłowski (2019) documents such destruction in the ten-month-long transition period. Yet the pace of the transition factors in our model in another way also: by reducing the  $\epsilon$  or uncertainty associated with the transformation. One could make the argument that a long protracted transition is less uncertain than a revolutionary turnover. In light of this, countries with roundtable-style, yet shorter transformations would be more uncertain (represented with higher  $\epsilon$ ) and those with abrupt “velvet revolutions” would be most uncertain (highest *epsilon*) implying the lowest likelihood of selective purges, all things considered.

Returning to our opening example from, we can ask for our model’s implications for Ukraine: Did the Ukrainians in 2016 behave rationally by sending the Berkut unit to the

---

<sup>17</sup> Although roundtable-style negotiations took place in Hungary, Bulgaria, and even Czechoslovakia, they were much more fast-paced.

Eastern front to fight Russian separatists in Donetsk and Luhansk? The firing of Berkut and the appointment of new patrol police bears resemblance to a thorough purge (even though the Berkut continued to wear uniforms they were now part of the military dealing with separatists instead of protesting citizens). In light of our model, a thorough purge is the preferred choice when the costs of selection are too high relative to the low competence and extreme ideological nature of the officers. Indeed, while the Berkut agents were considered “elite”, their elite nature lay in their ability to beat people up, a skill in high demand on the Eastern front, but in low demand in downtown Kiiiv (at least prior to 2022).<sup>18</sup>

The underlying problem modeled here— the combination of personnel shortages with acute uncertainty concerning competence and loyalty of state employees to the former authorities — is ubiquitous. Even transitions devoid of any negotiations have to make personnel decisions, at least until new competent agents can be trained. Universally, these new democracies need employees who can choose democratic policy in a way that will match the “true state of the world.” Hence, the implications of our theoretical model extend well beyond the Polish context.

## References

- Bates, G., Cinar, I. and Nalepa, M. (2020), ‘Accountability by numbers: A new global transitional justice dataset (1946–2016)’, *Perspectives on Politics* **18**(1), 161–184.
- Bates, G. and Cole, G. (2021), Racial justice as transitional justice: Framing and public support for racial justice initiatives in the united states, *in* ‘Law and Society Conference in Chicago, May 27, 2021’.

---

<sup>18</sup> This was arguably a problem with many of the so-called “elite” units of law enforcement. According to the Polish daily *Gazeta Wyborcza*, the understanding of what it meant to have a desirable skill set was highly ambiguous in the verification aftermath: “citizens are not to be beaten, even when they cannot defend themselves and even after night fall when it is harder to document.” (Wyborcza 1989)

- Callander, S. et al. (2008), ‘A theory of policy expertise’, *Quarterly Journal of Political Science* **3**(2), 123–140.
- Chincinski, T. (2002), ‘Bydgoski marzec 1981 roku’, *Biuletyn Instytutu Pamięci Narodowej* **2**(12 (23)).
- Crabtree, C., Kern, H. L. and Siegel, D. A. (2020), ‘Cults of personality, preference falsification, and the dictator’s dilemma’, *Journal of Theoretical Politics* **32**(3), 409–434.
- Egorov, G. and Sonin, K. (2011), ‘Dictators and their viziers: Endogenizing the loyalty–competence trade-off’, *Journal of the European Economic Association* **9**(5), 903–930.
- Elster, J., Offe, C., Preuss, U. K., Boenker, F., Goetting, U. and Rueb, F. W. (1998), *Institutional design in post-communist societies: Rebuilding the ship at sea*, Cambridge University Press.
- Epstein, D. and O’halloran, S. (1999), *Delegating powers: A transaction cost politics approach to policy making under separate powers*, Cambridge University Press.
- Gehlbach, S. (2021), *Formal models of domestic politics*, Cambridge University Press.
- Gerring, J. (2016), *Case study research: Principles and practices*, Cambridge university press.
- Gliwa, M. (n.d.), ‘Region rzeszowski nszz solidarność“ po 13 xii 1981 r.’.  
**URL:** <https://polskiemiesiace.ipn.gov.pl/mie/wszystkie-wydarzenia/grudzien-1981/historia/mapa-regionow-nszz-soli/116681,Region-Rzeszowski-NSZZ-Solidarnosc-po-13-XII-1981-r.html>
- Główny Urząd Statystyczny, P. (1988), ‘Mały rocznik statystyczny’, *Mały Roczniczek Statystyczny* .
- Główny Urząd Statystyczny, P. (1985), ‘Mały rocznik statystyczny’, *Mały Roczniczek Statystyczny* .

- Greitens, S. C. (2016), *Dictators and their secret police: Coercive institutions and state violence*, Cambridge University Press.
- Grzymala-Busse, A. M. (2002), *Redeeming the communist past: The regeneration of communist parties in East Central Europe*, Cambridge University Press.
- Huber, J. D. and Shipan, C. R. (2002), *Deliberate discretion?: The institutional foundations of bureaucratic autonomy*, Cambridge University Press.
- Instytut Pamięci Narodowej, . (2021), ‘Elektroniczny inwentarz archiwalny akt spraw karnych z okresu stanu wojennego’.
- URL:** <https://polskiemiesiace.ipn.gov.pl/mie/wszystkie-wydarzenia/grudzien-1981/represje/sprawy-karne/119225,Sprawy-karne.html>
- Kaminski, L. and Waligora, G. (2010), ‘Nszz solidarnosc 1980–1989’.
- Komar, A. B. and Niedzialek, M. C. (1990), ‘Przyczyny zwolnien z resortu spraw wewnetrznych’, *Humanizacja Pracy* **6**(1), 40–52.
- Kozłowski, T. (2019), *Koniec imperium MSW: transformacja organow bezpieczenstwa panstwa 1989-1990*, Instytut Pamięci Narodowej–Komisja Scigania Zbrodni przeciwko Narodowi.
- Lorentzen, P., Fravel, M. T. and Paine, J. (2017), ‘Qualitative investigation of theoretical models: the value of process tracing’, *Journal of Theoretical Politics* **29**(3), 467–491.
- Mack, E. (2018), Robert Nozick’s Political Philosophy, in E. N. Zalta, ed., ‘The Stanford Encyclopedia of Philosophy’, Summer 2018 edn, Metaphysics Research Lab, Stanford University.
- Marat, E. (2018), *The Politics of Police Reform: Society Against the State in Post-Soviet Countries*, Oxford University Press.

- McCarthy, L. A. (2015), *Trafficking Justice: How Russian Police Enforce New Laws, from Crime to Courtroom*, Cornell University Press.
- Mill, J. S. (1869), *A System of Logic, Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation*, Harper and brothers.
- Ministerstwo Spraw Wewnętrznych, . (1971-1990), Ministerstwo spraw wewnętrznych, jednostki organizacyjne. funkcjonariusze, pracownicy cywilni. zatrudnienie, ruch kadr w latach 1971-1990. raporty, zestawienia, notatki, informacyjne, korespondencja.
- Montagnes, B. P. and Wolton, S. (2019), ‘Mass purges: Top-down accountability in autocracy’, *American Political Science Review* **113**(4), 1045–1059.
- Nalepa, M. (2022), *After Authoritarianism: Transitional Justice and Democratic Stability, Political Economy of Institutions and Decisions*, Cambridge University Press.
- Nalepa, M. and Pop-Eleches, G. (2022), ‘Authoritarian infiltration of organizations: causes and consequences’, *The Journal of Politics* **84**(2), 000–000.
- Nozick, R. (1978), *Anarchy state and utopia*, Blackwell, London, England.
- Oseka, P. (2008), ‘Funkcjonariusze sb 1970-1989—drogi awansu i modele kariery’, *Wykluczeni* (1), 108–124.
- Paine, J. (2022), ‘Reframing the guardianship dilemma: How the military’s dual disloyalty options imperil dictators’, *American Political Science Review* doi:10.1017/S0003055422000089, 1–18.
- Piotrowska, B. M. (2020), ‘The price of collaboration: how authoritarian states retain control’, *Comparative Political Studies* **53**(13), 2091–2117.
- Piotrowski, P. (2003), ‘Struktury Sluzby Bezpieczenstwa MSW 1975-1990’, *Pamiec i Sprawiedliwosc* **3**(1), 51–108.

- Popova, M. (2012), *Politicized justice in emerging democracies: a study of courts in Russia and Ukraine*, Cambridge University Press.
- Przeworski, A. and Teune, H. (1981), 'The logic of comparative social inquiry, rev. ed', NY: *Wiley & Sons*.
- Putnam, R. D., Leonardi, R. and Nanetti, R. Y. (1994), Making democracy work, in 'Making democracy work', Princeton university press.
- Redakcja (2006), 'Wspomnienia po latach'.  
**URL:** <https://dziennikpolski24.pl/wspomnienia-po-latach/ar/1726036>
- Ritter, E. H. and Conrad, C. R. (2016), 'Preventing and responding to dissent: The observational challenges of explaining strategic repression', *American Political Science Review* **110**(1), 85–99.
- Scharpf, A. and Glasel, C. (2020), 'Why underachievers dominate secret police organizations: evidence from autocratic argentina', *American Journal of Political Science* **64**(4), 791–806.
- Skocpol, T. (1979), *States and social revolutions: A comparative analysis of France, Russia and China*, Cambridge University Press.
- Tarrow, S. (2010), 'The strategy of paired comparison: toward a theory of practice', *Comparative political studies* **43**(2), 230–259.
- Woldense, J. (2018), 'The ruler's game of musical chairs: Shuffling during the reign of ethiopia's last emperor', *Social Networks* **52**, 154–166.
- Wyborcza, G. (1989), 'Jeszcze w sprawie pobicia'.  
**URL:** <https://classic.wyborcza.pl/archiwumGW/6004913/Jeszcze-w-sprawie-pobicia>
- Zakharov, A. V. (2016), 'The loyalty-competence trade-off in dictatorships and outside options for subordinates', *The Journal of Politics* **78**(2), 457–466.

in the June 4, 1989 elections. Solidarity's victory in those elections was the beginning of the end of communism in Europe.

Reforming the *Sluzba Bezpieczenstwa* (SB), as the communist security services were called became a contentious topic in the Roundtable negotiations. Its sheer size made dealing with the SB challenging. In 1989, SB employed 24,308 officers (Kozlowski 2019). With a long history of repressing dissidents, SB was perceived as a key pillar of authoritarianism, but any staffing policy had to respond to challenges facing the new democratic state. Poland was in dire need of a screening mechanism.

Poland's dual regime change described above—from authoritarianism to democracy and from market socialism to capitalism—meant that fragile democratic institutions exacerbated the threat as large-scale property crimes and corruption brought about by the privatization of formerly state-owned enterprises. Poland needed competent officers trained in fighting corruption and white-collar crime.

To a large extent, the former SB agents were better positioned to combat white-collar crime, from corruption to industrial espionage, than insufficiently trained outsiders. Hence, the new security system was bound to include at least some personnel trained under the authoritarian SB.

With this in mind, the first democratically sanctioned cabinet appointed in July 1990 49 regional “verification” commissions(Kozlowski 2019) (RVC). Even though all officers employed by SB at the end of July 1989 were fired, each of them could apply for a position in the newly created Office for National Protection (*Urzad Ochrony Panstwa, UOP*) after satisfying two conditions:

1. they had to be positively verified by the relevant (regional) commission;
2. they had to be younger than 55

Positive verification was not synonymous with re-employment in the new secret service: While some 14,000 officers participated in the verification process, there were only c.a. 5,000

vacancies in UOP. Each RVC included an UOP representative, the head of local police, a universally trusted local activist as well as MPs, and senators. This composition was more or less uniform in each *województwo* (as the 49 regions in Poland are called).

With little more than a few weeks to review hundreds of personnel files, RVC resources were thinly stretched. Given how little guidance came from the center, the variation seen in the verification outcomes in the introduction to this article might not seem surprising.

Perhaps some RVCs only approved officers who achieved socially desirable outcomes, while rejecting all remaining <sup>7</sup>; perhaps others rejected only those who were proven guilty of transgressions, while retaining other candidates.<sup>8</sup>

Though such idiosyncratic explanations are consistent with high variation we observe, our model provides a theoretically consistent reason for the variation.

A sub-national study like ours offers a rare opportunity to control for broader country-specific variables. Moreover given that the 49 RVCs operated fairly independently from one another in making the initial verification decision, this kind of analysis is reminiscent of subnational analysis of, say, state legislatures in the 50 United States.

This is crucial in light of many criticisms leveraged against the paired comparisons approach. Tarrow (2010). for instance, points out that the first instances of using this method by Skocpol (1979) and Putnam et al. (1994) paired cases that were distant either across time, space, or both. Tarrow argues that this makes it impossible for the analysis to pick up on hidden cultural or historical determinants that are responsible for the phenomenon in question but not accounted for by the independent variables of these authors.

---

<sup>7</sup> For instance, Kozłowski (2019) reports that in Tarnobrzeg 80 % of the applicants were verified negatively, as a positive evaluations were only offered to those who could provide evidence of contributing a desirable outcome from the point of view of dissidents

<sup>8</sup> Also Kozłowski (2019) describes that in Gdansk, the stronghold of the Solidarity trade union, RVC members only opined negatively case with concrete evidence of harm inflicted on dissidents (p.206).

## 4.1 Data

Ultimately, out of the 14034 security officers that underwent verification, 10439 (or 74%) were verified positively and could go on to apply to work in the new democratic security service. The range of officers whose files the RVCs analyzed and found unqualified to serve in the new democratic security services (negatively verified) runs from 7% to 79%, allowing us to see all three types of purges considered in the model within one country.

To shed light on the determinants of the type of purges, we analyze a number of historical and archival sources, compiling them into an original dataset.

The archival data from files of the RVCs (*Wojewodzkie Komisje Kwalifikacyjne*) contain, for each of the 49 wojewodztwa, information on the severity of purges, our key variable of interest. The files allow us to extract:

1. the number of officers that have undergone verification;
2. the number of those verified positively and negatively;
3. the time each RVC spent on deliberation;
4. the proportion of officers that appealed the negative decision of the commission;
5. the outcome of the appeal

An example of a page from the files constituting our data is provided in Figure 5.

However, rather than focusing on the proportion of negatively verified officers, the empirical interpretation of our model ought to focus on the type of purges conducted in each wojewodztwo. The type of purge is best captured as a combination of the verification outcome *and* the amount of time that each RVC put into adjudicating the individual cases. Since selective purges require careful reading of files and interviewing candidates, we expect them to be the most time-consuming.

Hence, we operationalize a *selective purge* as one where the RVC took a lot of time deliberating each case; a *thorough purge* as an instance where RVCs which took the least

# A Appendix

## A.1 Formal Model

First we derive baseline model conditions ensuring the optimality of choosing the selective purge over both the thorough purge and “no purge.” This requires solving simultaneously the pair of inequalities:

$$\begin{cases} -\epsilon < -\epsilon(1 - \theta\mu) - \theta\mu x_M - v \\ -\epsilon(\theta(1 - \mu)) - \theta\mu x_M - x_E(1 - \theta) < -\epsilon(1 - \theta\mu) - \theta\mu x_M - v \end{cases}$$

The top inequality, corresponding to when a selective purge is better than a thorough purge reduces to:

$$-\epsilon < -\epsilon + \epsilon\theta\mu - \theta\mu x_M - v$$

and eventually to

$$v < \mu\theta(\epsilon - x_M)$$

The second inequality, corresponding to when a selective purge is better than “no purge” expands to

$$-\epsilon\theta + \epsilon\theta\mu - \theta\mu x_M - x_E + x_E\theta < -\epsilon + \theta\mu\epsilon - \theta\mu x_M - v$$

and further to

$$\epsilon(1 - \theta) - x_E(1 - \theta) + v < \mu(\theta\epsilon - \theta x_M - \epsilon\theta + \theta x_M)$$

, which further reduces to

$$v < (x_E - \epsilon)(1 - \theta)$$

Jointly these two inequalities make up the constraints that define  $v^*$ , the constraint that

the cost of conducting a selective purge must satisfy.

### A.1.1 Alternative specification of preferences

We can also calculate the utilities to the regional verification commission from a selective, a thorough and no purge assuming, as is customary in models of political economy, using quadratic preferences of all players:

$$u_A(p; a_P) = -(p + \omega - x_T)^2$$

, where  $T \in \{M, E\}$

$$u_P(p; a_P) = -(p + \omega)^2 - v * \mathbb{1}_{sp}$$

Although this does not change the main findings of the baseline model, it will be useful in the extension to appeals, as it will guarantee the uniqueness of the results. Quadratic preferences will ensure that the best response of selecting  $p$  equal to the ideal point of the incompetent or replacement agent implementing policy is unique.

It will affect the payoffs to the regional verification commission in the following way. The utility from a thorough purge will be given by

$$U_P(p^*; nocost, tp) = -(\epsilon)^2$$

This is a straightforward consequence of the fact that a thorough purge results in the appointment of the replacement agent who plays his optimal strategy described above as  $p^* = 0$ .

The utility from a selective purge is given by:

$$\begin{aligned}
U_P(p^*, \text{paycost}, sp) = & \underbrace{-\theta\mu(x_M)^2}_{\text{moderate, competent}} \\
& \underbrace{-(1-\mu)(\epsilon)^2}_{\text{replaced any incompetent}} \\
& \underbrace{-(1-\theta)\mu(\epsilon)^2}_{\text{replaces extreme competent}} - v
\end{aligned}$$

And the utility from no purge is given by:

$$\begin{aligned}
U_P(p^*, \text{nocost}, np) = & \underbrace{-\theta\mu x_M^2}_{\text{moderate, competent}} \\
& \underbrace{-(1-\theta)\mu x_E^2}_{\text{extreme, competent}} \\
& \underbrace{-\theta(1-\mu)(0-(x_M+\epsilon))^2 * \frac{1}{2} + (0-(x_M-\epsilon))^2 * \frac{1}{2}}_{\text{moderate incompetent}} \\
& \underbrace{-(1-\theta)(1-\mu)(0-(x_E+\epsilon))^2 * \frac{1}{2} + (0-(x_E-\epsilon))^2 * \frac{1}{2}}_{\text{extreme incompetent}}
\end{aligned}$$

After simplifying, the expected payoffs to the Regional Verification Commission from no and selective purges, respectfully are

$$u_P(np) = -\epsilon^2(1-\mu) - x_E^2(1-\theta) - \theta x_M^2$$

and

$$u_P(sp) = -x_M^2\theta\mu - \epsilon^2(1-\theta\mu) - v$$

The equivalent of Proposition 1 in terms of these quadratic preferences is that for selective purge to constitute the optimal choice of the New Democratic Politician,  $v$  must be smaller than the minimum of  $\{\theta\mu(\epsilon^2 - x_M^2), x_M^2\theta(1-\mu) + (1-\theta)(x_E^2 - \epsilon^2\mu)\}$ .

### A.1.2 Alternative specification of the main result

An alternative way of presenting our results is to consider when a thorough purge is better than a selective purge and no purge. We already know this happens when the cost of con  $v$  is high.

To explore this more systematically, we can derive formally the conditions under which a thorough purge is preferred. Based on (2) and (1), the New Democratic Politician should choose a thorough purge when

$$\begin{cases} -\epsilon > -\epsilon(\theta(1 - \mu)) - \theta\mu x_M - x_E(1 - \theta) \\ -\epsilon > -\epsilon(1 - \theta\mu) - \theta\mu x_M - v \end{cases}$$

, which reduces (see appendix for detailed calculations) to  $\mu < \min\left\{\frac{(1-\theta)(x_E-\epsilon)}{\theta(\epsilon-x_M)}, \frac{v}{\theta(\epsilon-x_M)}\right\}$ .

This result is summarized in Proposition 2

**Proposition 2.** *The verification commission chooses a thorough purge over “no purge” or selective purge if and only if the proportion of competent agents is ,  $\mu$  is sufficiently small, that is,  $\mu < \mu^*$ , where  $\mu^* = \min\left\{\frac{(1-\theta)(x_E-\epsilon)}{\theta(\epsilon-x_M)}, \frac{v}{\theta(\epsilon-x_M)}\right\}$ .*

This result is consistent with the analysis above, indicating that an increase in the proportion of extreme agents makes a thorough purge more attractive (and a selective purge less attractive). Given the shared denominator which of these two constraints is binding depends on the magnitude of  $v$  and specifically on whether  $v > v^* = (1 - \theta)(x_E - \epsilon)$ . Intuitively, for low costs of verification, the constraint associated with a selective purge is binding, whereas for high costs of verification, the constraint associated with “no purge” binds. When no purge binds, the effect of the policy shock,  $\epsilon$  is critical.

When the shock of policy implementation is low, competent agents pose a liability because they are effective at absorbing the shock and forcing making the final policy outcome coincide with their ideal point. In this event, the verification would rather roll the dice and deal with  $\epsilon$ . Yet, as we argue, especially below, in volatile times of transition, policy

shocks are typically at least moderate, hence in Figure 3 below, where we clarify the effects of the parameters of interest on the incentive to conduct selective purge, we assume that  $\epsilon$  is at least .4.

## A.2 Full empirical results

The figure below places the case of Ukrainian purges in the context of other European former autocracies and post-conflict countries, using the Global Transitional Justice Database as a source. We can see that Ukraine is not isolated in its decision to refrain from purging the enforcement apparatus (in fact, given Albania’s and Romania’s choices, it was not even unique in concentrating purges on the bureaucracy while exempting law enforcement). Yet it was the country in the region where the differential between purges in the bureaucracy and law enforcement was highest. The case we use in this article—Poland, on Figure 9 uses data from the Global Transitional Justice Dataset Nalepa (2022), which disaggregates post-authoritarian purges into purges of law enforcement and purges of the bureaucracy. As can be seen, Ukraine up until 2014 had practically no attempts at purging its law enforcement apparatus.<sup>19</sup> According to Erica Marat (2018), even after Euromaidan Ukraine was slow to embark on a purge of the enforcement apparatus, in part because the war in Eastern Ukraine offered a convenient transfer point of the most militarized police forces. This is corroborated the visualization below in Figure 10 that zooms in on Ukraine’s enforcement apparatus purge trajectory and is also based on the GTJD.

### A.2.1 Alternative definition of purges

To create an alternative definition of purges, we use the median values of the two key variables (proportion verified negatively and hours per application) to identify the high

---

<sup>19</sup> What the Global Transitional Justice Dataset (GTJD) considers an event is something as minor as putting purge proposal on the legislative agenda, even if it were to fail the next day. Hence what these data indicate is that even by this very conservative measure, Ukraine did nothing at all to demilitarize its authoritarian-era police force. Another figure placed in the Empirical Appendix places Ukraine in the context of other European countries and additionally compares purges of law enforcement with purges of the bureaucracy.

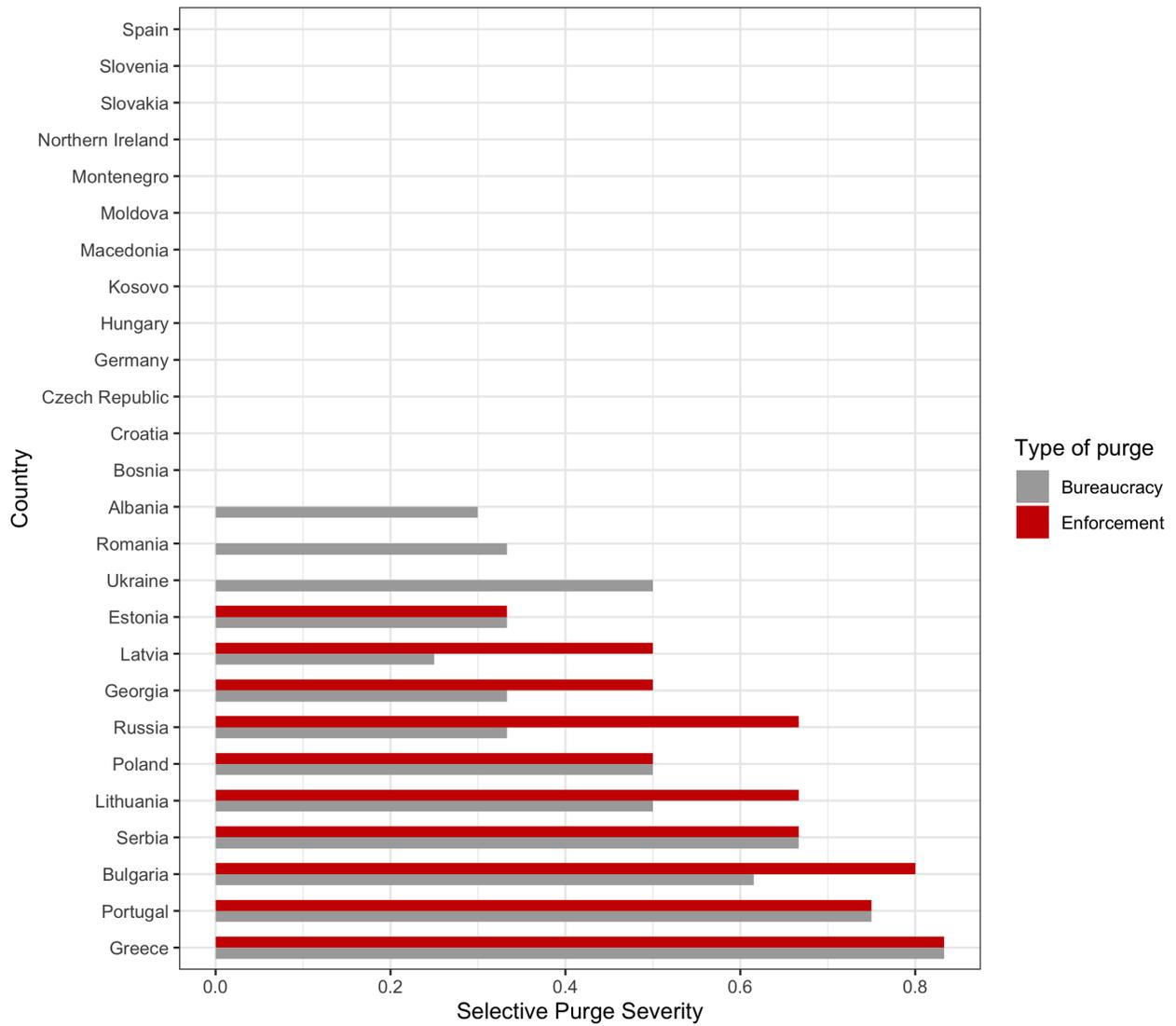


FIGURE 9: *Purges of Law Enforcement in Post-communist Europe 1990-2016* (Source: GTJD Nalepa (2022))

TABLE 6: *Full data on the key variables*

Wojewodztwo	Repression	Population density	Competence	Purge	Purge (median)
Białostockie	0.473	66.793	0.232	selective	selective
Białkopodlaskie		55.703	0.348	selective	selective
Bielskie	0.694	235.853	0.331	thorough	thorough
Bydgoskie	0.541	104.725	0.363	selective	selective
Chełmskie	0.621	63.192	0.366	selective	selective
Ciechanowskie	1.000	65.718	0.193	no	selective
Częstochowskie	0.471	124.976	0.270	thorough	thorough
Elbląskie	0.364	76.471	0.324		
Gdańskie	0.346	189.546	0.346		
Gorzowskie	0.976	56.872	0.272	selective	selective
Jeleniogórskie	1.000	117.493	0.254	thorough	thorough
Kaliskie	0.565	106.941	0.223	thorough	selective
Katowickie	0.723	588.932	0.177	thorough	thorough
Kieleckie	0.853	127.120	0.172	thorough	thorough
Konińskie	0.294	89.375	0.423	no	no
Kozalińskie	0.587	57.828	0.225	selective	selective
Krakowskie	0.586	371.635	0.278		
Krośnińskie	0.313	83.339	0.230	selective	selective
Legnickie	0.886	125.167	0.294	selective	selective
Leszczyńskie	0.375	90.419	0.204	selective	selective
Łódzkie	0.692	748.753	0.411	no	no
Łomżyńskie	0.667	43.397	0.247	thorough	no
Lubelskie	0.715	148.145	0.153		
Nowosadeckie	1.000	119.692	0.142	selective	selective
Olsztyńskie	0.725	58.871	0.204	selective	selective
Opolskie	0.800	118.102		thorough	thorough
Ostroleckie		59.126	0.225	selective	selective
Piłskie	1.000	53.586	0.229		
Piotrkowskie	0.818	101.037	0.154	thorough	thorough
Płockie	0.800	100.078	0.365	no	selective
Poznańskie	0.291	159.244	0.326	selective	selective
Przemyskie	0.400	89.227	0.194	selective	selective
Radomskie	0.333	100.041	0.199	no	no
Rzeszowskie	0.709	157.221	0.213	thorough	thorough
Siedleckie	0.333	71.332	0.242	selective	selective
Sieradzkie	1.000	82.416	0.224	no	no
Skierniewickie	0.250	53.603	0.302	selective	selective
Słupskie	0.783	100.025	0.533	thorough	thorough
Suwałskie	1.000	42.803	0.478		
Szczecińskie	0.420	94.430	0.300		
Tarnobrzesckie	0.000	92.392	0.133	thorough	thorough
Tarnowskie	0.393	154.541	0.211	selective	selective
Toruńskie	0.504	119.783	0.204	no	no
Wałbrzyskie	1.000	177.063	0.309	no	selective
Warszawskie	0.539	636.800	0.247	no	no
Włocławskie	0.438	96.751	0.354	no	no
Wrocławskie	0.842	177.175	0.161	thorough	thorough
Zamojskie	1.000	69.900	0.228	no	no
Zielonogórskie	0.909	72.846	0.206	no	no

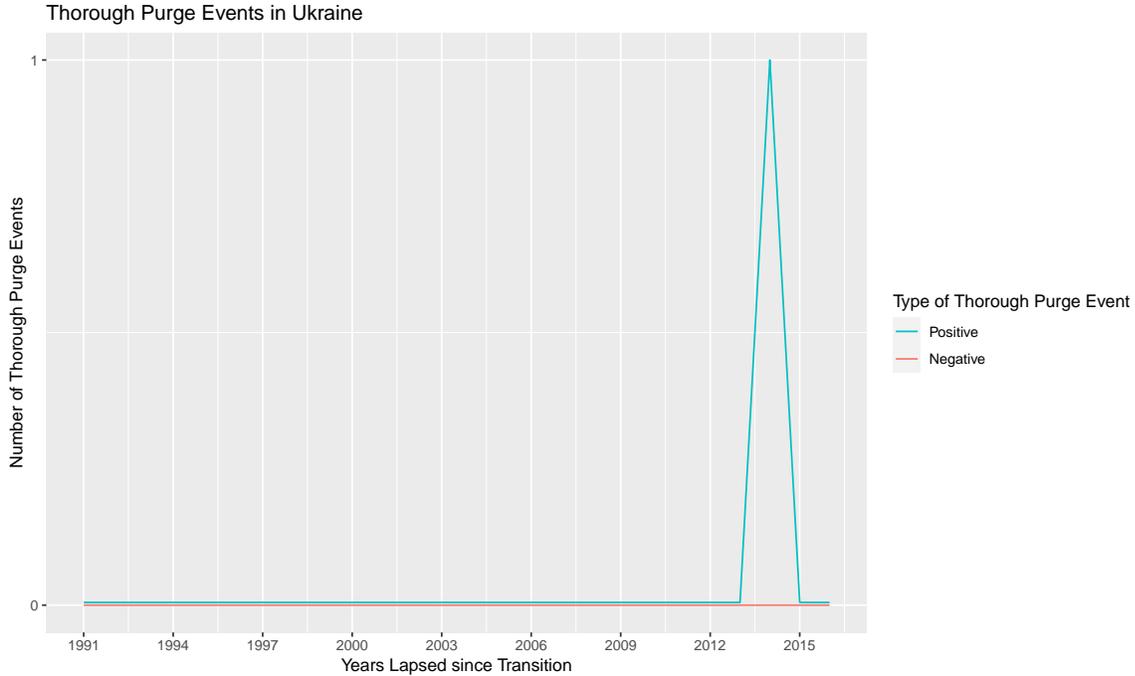
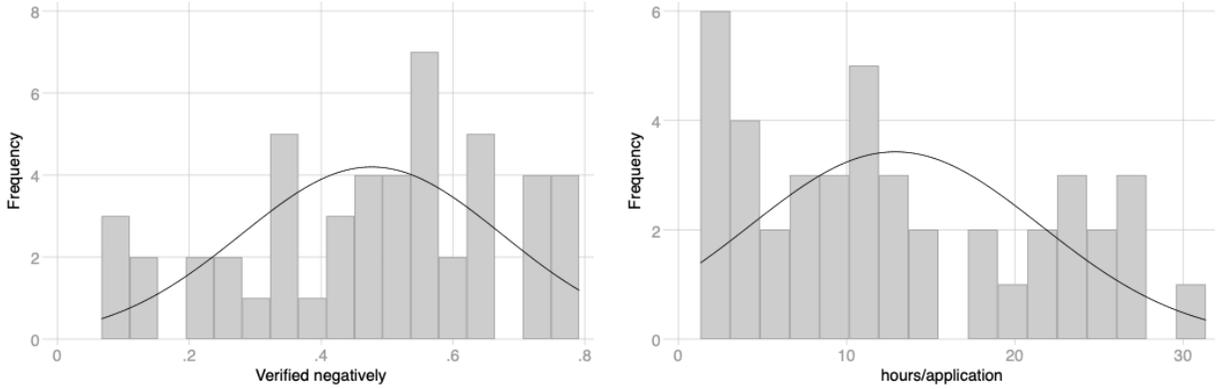


FIGURE 10: *Trend in Law Enforcement Purge Events in Ukraine 1991-2016, Source: Global Transitional Justice Dataset Bates et al. (2020)*

and low values of negative verification and time spent per application (Figure 11). Because the median time per application is lower than the mean (11.3, as compared to 12.9) the proportion of selective purges is higher. To the contrary the median proportion verified negatively is higher than the mean (0.51, compared to 0.48) so for the low time values more wojewodztwa are classified as having no purges (Figure 7).

TABLE 7: *Purge specification using median values*

Purge type	Proportion verified negatively	Hours per application	Frequency	Percent
No	<0.51	<11.3	10	23.8
Selective		>11.3	21	50.0
Thorough	>0.51	<11.3	11	26.2



(A) *Proportion of officers verified negatively*

(B) *Hours spent per application*

FIGURE 11: *Histograms of the variables key for purge categorization*

### A.2.2 Paired comparison

For the Most Similar System Design to isolate the effect of the key independent variable the chosen cases need to be as similar as possible on the relevant control variables. In the text, we describe the key similarities in terms of the model-derived variables, as well as the cultural background of the chosen pairs of wojewodztwa. However, the two sets of areas are also similar on other variables reported in Table 8.

Most strikingly, wojewodztwa within the two pairs are nearly identical in terms of the proportion of migrants. This variable captures the extent of the post-World War II population re-settlement (Nalepa and Pop-Eleches 2022) and could affect the type of purges by influencing the extent to which people within the wojewodztwa were part of networks lasting for many generations. Hence, our analysis shows that contemporary population density, rather than long-term effects of re-settlement affected the type of purges used. Similarly, the scale of opposition (Density Solidarity) and security penetration (Density SB) within the two pairs were comparable. Hence, it was the intensity of repression, rather than the mere presence of the opposition or security apparatus that affected the types of purges used. Finally, the wojewodztwa, and in particular bydgoskie and wloclawskie, had a very similar economic make-up, addressing questions of potential endogeneity of competence.

Finally, there are two robustness checks to note. First, our analysis is robust to alternative measures of extremeness. As a reminder, our main measure considers those sentenced as a proportion of individuals tried under the Martial Law 1981-1983. The alternative measure looks at the cases and sentences as a proportion of Solidarity members in each wojewodztwo. These alternative numbers co-vary in line with the theory. Second, the pairs of wojewodztwa are classified in the same way, regardless of whether we use mean or median values for purge classification.

TABLE 8: *Additional control variables*

	Tarnowskie	Rzeszowskie	Bydgoskie	Wlclawskie	Mean	SD	Min	Max	Source
Repression	0.3900	0.7100	0.5409	0.4375	0.0002	0.0002	0.0000	0.0006	(1)
Cases/Solidarity	0.0002	0.0004	0.0005	0.0002	0.0003	0.0003	0.0000	0.0011	(1)(2)
Sentences/Solidarity	0.0001	0.0003	0.0003	0.0001	0.0002	0.0002	0.0000	0.0006	(1)(2)
Population density	154.50	157.20	104.73	96.751	140.74	145.91	42.803	748.75	(3)
Competence	0.2114	0.2134	0.3631	0.3542	0.2650	0.0881	0.1329	0.5333	(4)
Purge (median)	selective	thorough	selective	no					
Officers verified	139.00	206.00	284.00	164.00	223.69	133.53	104.00	786.00	(5)
Density Solidarity	0.2431	0.3472	0.2713	0.2113	0.2388	0.0822	0.0847	0.4439	(2)(3)
Vacancies	14.200	19.700	16.800	12.100	15.790	4.406	5.7000	31.400	(6)
Migrants	0.0273	0.0255	0.0529	0.0538	0.2848	0.3541	0.0000	1.0000	(7)
Density SB	0.0004	0.0005	0.0004	0.0005	0.0005	0.0001	0.0003	0.0008	(3)(4)
Production heavy	11.500	36.900	29.700	20.300	24.216	10.725	4.0000	51.200	(8)
Production light	2.2000	7.0000	6.1000	4.6000	15.290	12.366	1.9000	48.500	(8)
Production produce	14.000	34.500	20.200	22.200	23.680	14.095	3.5000	59.000	(8)

Data sources:

- (1) Instytut Pamięci Narodowej (2021)
- (2) Kaminski and Waligora (2010); Nalepa and Pop-Eleches (2022)
- (3) Główny Urząd Statystyczny (1985); Nalepa and Pop-Eleches (2022)
- (4) Piotrowski (2003)
- (5) IPN Archival files of the RVCs: BU/3546/6 - BU/3546/54
- (6) Ministerstwo Spraw Wewnętrznych (1971-1990)
- (7) Nalepa and Pop-Eleches (2022)
- (8) Główny Urząd Statystyczny (1988); Nalepa and Pop-Eleches (2022)