Justice or Practicality: the operation of vetting commissions in the process of purging the enforcement apparatus in Poland

Monika Nalepa[1] and Barbara Piotrowska[2]

[1]Department of Political Science,The University of Chicago, Chicago, IL 60615, USA
[2]Department of Political Economy, King's College London, London, UK

November 4, 2021

Abstract

The aftermath of the murder of George Floyd, Breonna Taylor, and others in 2020, brought numerous calls for de-funding the police countered with more moderate proposals for just "throwing out the bad apples." Although the dilemma facing the US seems unique, numerous new democracies recovering from authoritarian rule have had to reckon with members of their repressive and enforcement apparatuses. One solution was to just disband enforcement agencies in their entirety and build new ones, along with institutions of the new democratic state. Another, competing approach, was to pick out and only fire the "bad apples." The latter approach is one a new democracy may want to take when it is strapped for competent enforcement agents and when competent —as opposed to predominantly loyal—agents were indeed employed by the preceding autocratic regime. This paper analyzes the costs and benefits of conducting, what we call, a thorough purge—understood as dismissing everyone in a certain law enforcement agency—and a selective purge, where the competence and loyalty of agents is evaluated on a case-by-case basis. We propose a theory that formalizes the loyalty-competency trade-off facing new democracies deliberating how to reform their security apparatus. Our model allows to make predictions about the causes as well as consequences of thorough as opposed to selective purges. We corroborate our theory using newly acquired data from the Institute of National Remembrance (IPN) in Poland on the operation of verification commissions in 49 sub-national regions in 1990.

1 Introduction: disband or reform?

Protection of the rights of life, liberty, property, and contract is a fundamental function of even the most minimal conception of the state (Nozick 1978; Mack 2018). Hence, it is crucial for governments to ensure that the institutions offering protection, such as the police, security services or courts, function properly. Yet, this is not always the case. In 2020, hundreds of thousands of protesters took to the American streets in response to the brutal murder of George Floyd and Breonna Taylor at the hands of uniformed policemen. As opposed to the local protests in the aftermath of similar incidents of police brutality in the past, this time the postulates went beyond just demanding a fair trial for the perpetrators directly involved. Some members of the Black Lives Matter movement demanded structural change. The calls went as far as "de-funding the police," interpreted as disbanding entire police departments and forcing municipalities to start building police departments anew. Other supporters of the plight of Black Lives Matter (BLM), however, called for moderation. Concerned about what a complete lack of police forces might spell for inner city neighborhoods, they urged local leaders and policy makers to instead engage in reforming the existing enforcement apparatus.

The dilemma facing local policymakers in the US is not unique. 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 addition to this, most agencies had also been infiltrated by the KGB.

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 politicians wanted to limit the influence of old agencies entirely resulting in a thorough purge. On the other end of the spectrum were the pragmatists. They argued that law enforcement is but a tool in the hands of politicians and just as the agency once supported the communist authoritarian regime, it could now serve the needs of the new

Polish democracy. The secret service should be reformed, but in a way that retains the experts whose origins were in the communist secret police force" (Kozłowski 2019, p.289).

Not unlike the dilemma facing BLM post 2020, the choices before newly democratized states were to either disband the enforcement apparatus in its entirety or engage in careful verification of individual files in an attempt to "pick out the bad apples."

The meaning of "bad" in the term "bad apples" is ambiguous however. It could refer to enforcement agency officers who persecuted the anti-authoritarian opposition and "deserved" to be punished by members of the incoming democratic regime; or it could refer to incompetent agents, who were appointed to their positions due to their ideological loyalty to the communist leaders. The latter kind of officers are bad not because they persecuted the opposition, but because they are incompetent and retaining them offers no value to the new state.

In this paper, we use the case of Post-Communist Poland to illustrate how policy-makers resolve the dilemma of whom to fire and how their sense of practicality—the demand for experts skilled in fighting crime —keeps their desire for justice (and potentially revenge) in check. The next section places our question within the general literature on the loyalty competency trade-off. Here, we explain how the transitional justice context adds nuance to this classical and well-researched dilemma. The following section presents a simple theoretical model that systematizes criteria for determining when policy-makers are motivated by justice and when they are motivated by practicality. Section four makes the argument for why Poland, with its 49 regions and 49 verification commissions, is the ideal case for studying our question. This section also introduced readers to the archival data from the Institute of National Remembrance. Section 5 provides preliminary results and concludes.

2 Loyalty, Competency and Justice

Originally, the loyalty-efficiency tradeoff was formalized by Egorov and Sonin (2011) who used it to explain why rulers, but particularly dictators, have to balance their desirability for loyal agents on the one hand and 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.

Alexandre Debs (2016), although he does not talk specifically about the incentives of leaders and their rank-and-file, adds nuance to this point of view by drawing attention to regime type. Specifically, uniformed personnel from former military regimes fair better under new democratic rulers than non-democratic rulers. Since democracies select rulers based on citizenship support and not brute force, military members do not pose a threat to new democracy in the same way they pose a threat to an autocracy. Debs uses this insight to draw conclusions about the speed and nature of regime change. He predicts that transitions from military rule to democracy should appear more swiftly than from other forms of dictatorship.

Other scholars have integrated social networks into analyzing the competency-loyalty tradeoff. For instance, Josef Woldense (2018) shows that rulers can alleviate it by shuffling employees laterally from one geographic region to another. Shuffling prevents 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.[1] Yet not posing a threat does not immediately imply loyalty. What if the agent just doesn't care to put in

[1] For a similar argument in the context of Putin's Russia, see Olimpieva (2021)

effort?

Jack Paine (2022) examines this aspect of the competency-loyalty tradeoff in the context of organizing security and repressive apparatuses by authoritarian rulers. In this interpretation of the tradeoff, 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.

In a final paper we discuss here on the loyalty-competency dilemma, Alexei Zakharov (2016) considers an 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.

Our model departs from Paine, Zakharov, Egorov and Sonin, and Woldense in several respects. Key among them is that we are less interested in the former authoritarian ruler and more in the incentives of democratic rulers who succeed him. In the democratic context, the trade-off between loyalty and expertise becomes nuanced by the fact that agents with expertise and weaker loyalty to the outgoing, authoritarian regime, may in fact be more

loyal to the new democratic regime. Hence, in some circumstance, the loyalty-competence trade-off can be averted. Moreover, since the new democrats inherit a state staff potentially loyal to the autocrat, their decision is better framed as one not about whom to appoint, but about 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 theory setting, is the decision to fire agents based on their repressive behavior towards dissidents. In this article, we treat this tendency towards repressive behavior to be independent of loyalty and competence.

3 The Model

In order to model the tasks before a policymaker in a new democracy deciding how to deal with members of an authoritarian security apparatus, we follow Nalepa (2021) in using the work horse 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 flipside of the delegation problem. Instead of choosing to delegate, the new democratic Politician chooses whether to retain an existing agent that was appointed by a different 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. The players are a new democratic policy maker with ideal point 0 in a uni-dimensional issue space and a former law enforcement agent who may be one of two types: (1) hired by the former authoritarian regime for loyalty or (2) hired by the former authoritarian regime for competence. The agent of the law enforcement apparatus also has an ideal point in the policy implementation space, which is unknown to the new democratic policy maker, x_t, where $t \in T = \{m, e\}$ With probability $\theta \in (0, 1)$ the agent is moderate (with ideal point at x_m); with probability

$(1-\theta)$, the agent is extreme (with ideal point at x_e).

We also assume that competent agents are better at detecting signals/implement policy with less noise. 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 ω with perfect precision, as in the standard delegation model. With probability $\mu \in (0,1)$ the agent is competent (can observe ω) and with probability $1-\mu$ he was appointed for loyalty to the autocrat. For now, we assume that θ and μ are independent, but this can be relaxed later. Assume that $0 > x_m < \epsilon m < x_e < 1$.

The democratic policy maker can pay a cost and uncover if the law enforcement agent was hired for loyalty or for competence (this is the individual level, selective purge). The cost will be later interpreted as investing in a verification commission that actually studies and evaluates the files.

For the baseline model, we analyze the following interaction between the policy maker, P and law enforcement agent, A.

In the first period, Nature chooses if the agent is competent and whether his ideal point is moderate or extreme. If one were to represent the former authoritarian ruler's ideal point at 1 an extreme agent is one who is more loyal to the former authoritarian ruler. In the second period, the policy maker observes the implemented policy (not having observed the noise) and chooses one of three actions $a_P \in \{tp, sp, np\}$:

- thorough purge (tp), fires everyone and appoints replacement who has the same ideal point as the policy maker, i.e., 0, but no competence (so cannot observe ω);
- selective purge (sp), where after paying cost v he can learn the competence (but not ideal point) of the agent and only fire the 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 (an untrained agent of the policy maker with identical ideal point, at 0, but no ability to learn ω) implements policy. The game ends and payoffs are collected.

Utility functions are defined by distance of implemented policy (so $p + \omega$) from players' ideal points. In the case of the policy maker, utility functions are also determined by the cost of appointing a verification commission (if one is appointed). The agents' utility is also shaped by the distance of implemented policy from ideal point, but also by the value of employment (not being purged).

The utilities are as follows:

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

, 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 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 through generalized backward induction (Kaminski 2019; Kaminski and Nalepa 2014).

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 $p^* = 0$.

Consider now the agent without competence. Like the replacement agent, he cannot

observe ω, 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 ω. Their maximization problem, in light of knowing ω leads them to chose

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

We can next move to the New Democratic Politician's A'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 slightly more complex expressions reflects what takes place when no purge takes place. Starting from the top, the moderate and competent former secret police agent will

adopt the policy that "absorbs" the shock ω, (which he observers) to arrive at his ideal point, x_M. There are θ (for moderate) * μ (for competent) such cases. Moving on to the incompetent agents, reflected in the next two lines, recall that these agents too strive to arrive at their ideal point but are unable to absorb the shock, ω 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:

$$-\epsilon(x_M(1-\mu)) - \theta\mu x_M - x_E(1-\theta) \tag{1}$$

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

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

To understand this expression, recall that the selective purge reveals complete information on both the competence and the ideology of the former agents. A plausible extension that follows assumes that agents have 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 will assume that this commission only gets observe the ideology, but not competence of the agents.

the above expression reduces to:

$$-\epsilon(1-\theta)\mu - \theta\mu x_M - v \tag{2}$$

What does the New Democratic Politician 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 thorough purge, the entirety of 1 is negative. However, the coefficient on ϵ is less than 1, so at least for some parameter values, 2 can be greater that $-\epsilon$.

Specifically, an ideal point of the moderate type of agent closer to the ideal point of the New Democratic Politician will 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 the uncertainty associated with policy implementation, then delegation is more likely). Another factor that will decrease 1 relative to $-\epsilon$ is the proportion of extreme agents $(1-\theta)$. Based on eyeballing the expression alone, it is not immediately clear what the effect of μ, the proportion of competent agents on the payoff from a thorough as opposed to no purge.

However a comparison of$-\epsilon$ with expressions 1 and 2 allows us to also formulate conditions under which a thorough purge will be more attractive than a selective purge or "no purge." The payoff associated with a thorough purge is flat: at $-\epsilon$ it does not depend at all on μ or θ. Conditions when a thorough purge is preferred over a selective or "no purge" can be are extracted from 2 and 1 as:

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

After simplifying the above expression we arrive at the following proposition

Proposition 1. *1. Assume that relative to the proportion of moderate agents, the magnitude of the uncertainty of policy implementation, is small (i.e., $\epsilon < \theta$. The New Democratic Proposition choose a thorough purge over "no purge" or selective purge if*

and only if the probability of competent agents is sufficiently large, specifically when $\mu > max\{\frac{\epsilon - v}{\epsilon(1-\theta)+\theta x_M}, \frac{\epsilon(1-x_M)-x_E(1-\theta)}{\theta x_M - \epsilon x_M)}\}$.

2. *Assume conversely, that relative to the proportion of moderate agents, the magnitude of the uncertainty of policy implementation, is is large (i.e.,* $\epsilon > \theta$*. The New Democratic Proposition choose a thorough purge over "no purge" or selective purge if and only if the probability of competent agents is moderate, that is, when* $\frac{\epsilon(1-x_M)-x_E(1-\theta)}{\theta x_M - \epsilon x_M)}\} > \mu > \frac{\epsilon - v}{\epsilon(1-\theta)+\theta x_M}$

Consider the first part of proposition 1. When the shock of policy implementation is low, competent agents pose a liability because they are effective at absorbing the shock making the final policy outcome coincide with their ideal point. In this event, the New Democratic Politician would rather roll the dice and deal with ϵ.

The second part of proposition 1, extends the logic of the tradeoff between competent agents and the policy shock of the first part. As the shock gets larger, New Democratic Politicians become willing to forgo thorough purges when the proportion of competent agents is high, though they will still choose a thorough over selective purge for moderate levels of competence on average.

Consider next an informal comparison of the payoff from a selective purge, summarized in 2 with the payoff from a thorough purge. We see here that increasing the cost of verification, v unambiguously decreases the value of the selective purge. But similarly, decreasing μ, the probability of the agent being competent increases the value of the selective purge relative to a thorough purge. This may be puzzling, but becomes clear upon deeper reflection: if few agents are competent, it may be worth paying the extra verification cost to find out who among them are the competent ones. Formally, the condition for a selective purge to be

preferred over a thorough or no purge, it has to be the case that:

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

After simplifying the above expression we arrive at the following proposition

Proposition 2. *The New Democratic Politician, P, chooses to pay a verification cost, v, and embark on a selective purge if and only if the probability of competent agents, μ, is sufficiently small, specifically when*

$\mu < min\{\frac{\epsilon - v}{\epsilon(1-\theta)+\theta x_M}, \frac{v - x_E(1-\theta) - \epsilon x_M}{\epsilon(\theta - x_M - 1)}\}$.

3.1 Discussion

To see the results of this analysis in yet a different way, the New Democratic Politician's decision problem can also be represented with the following two figures, showing the utility of the New Democratic Politician from his three separate actions as a function of μ for several fixed levels of ϵ (.4), θ(.2 top and .75 bottom), x_M (.1 top and .3 bottom), x_E (.45 top and .9 bottom) and v (.2 top and .75 bottom)

Figure 1 allows to illustrate the main propositions above. We see that selective purges provide the New Democratic Politician with the highest utility when the proportion of competent agents is lower, provided that the cost of verification is not too high. Perturbations of the parameters defining the ideal points of the former agents and even the proportion of moderate as opposed to extreme agents does change the optimality of selective purging as much as increasing the value of v does.

When v is high however, the horse race for the optimal purge action is between thorough purge (the dashed line) and no purge (the red lines). The upward sloping red line corresponds to a relatively low proportion of extreme agents whereas a the slightly downward sloping red line corresponds to the payoff from no purge when the proportion of extreme agents is high.

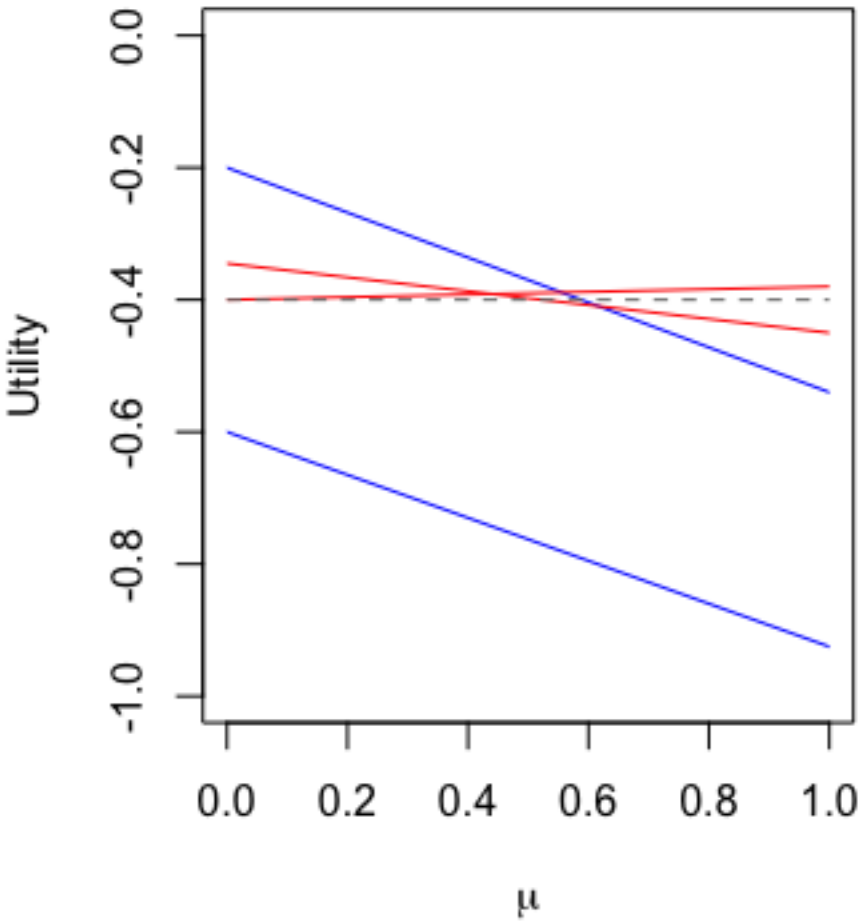


Figure 1: *New Democratic Politician's Utility from "selective purge" (blue) "no purge" (red) and thorough purge (dashed) as a function of μ, the proportion of competent former agents*
Note: Other parameters have been fixed at the following values: ϵ (.4), θ(.2 top and (.75 bottom), x_M (.1 top and .3 bottom), x_E (.45 top and .9 bottom) and v (.2 top and .75 bottom), where "red top" refers to the function increasing in μ

Consistently with Proposition 1, in the former case, a high level of competent agents urges the New Democratic Politician to abandon the idea of purging altogether, settling for the no purge choice. In the event that the proportion of competent agents is low, it is a toss up between "no purge" and "thorough purge", at least of the set of parameters we have chosen (the lines almost overlap). In contrast, in the latter case, when there is a high chance of extreme agents, the New Democratic Politician chooses Thorough Purge when the the level of competent agents is high, but no purge when the level is of competent agents is low. This is intuitive: as remarked, when most of the former agents are extreme *and* competent, they will be very effective at moving policy implementation closer to their ideal point, which is far from the New Democratic Politician's ideal point.

3.2 Extensions

In order to construct the baseline model described above, we had to make a number of simplifying assumptions, such as that the new democratic politician, after paying the cost of verification, uncovers the competence and loyalty of the new democratic politician with perfect precision and that once made, his decision to retain or fire a former agent of the *ancien régime* is final. Consequently, our model does not capture the right of rejected officers to appeal their decisions to a central verification commission, something that our archival research indicates as a common practice. In the first extension of our model, we will add two additional periods to our game. In the first, the officer, who knows he was fired but is not sure whether he was fired due to a thorough or selective purge can choose to reject the result and ask a central commission to reconsider the decision. The appeal comes at a cost to the ex-officer, but if his appeal is granted, he enjoys the benefit of office. The central commission, however, only learns the loyalty, not the competence of the appealing officer and bases its decision on that alone. Note, that this may lead the central commission to making mistakes and retaining a incompetent but moderate agent. The central commission is only motivated by its ideal point, which coincides with the ideal point of the new democratic

politician.

In a second extension, one could allow the loyalty and competence of candidates not to be independent as currently assumed, leading the central commission to infer competence from loyalty (if loyalty and competence are not independent, but correlated, loyalty effectively serves as a signal of competence).

A final extension that could be considered would relax the assumption of perfect selective purging. The New Democratic Politician could choose how much she wants to pay for the selective purge which would result in a imperfect signal of competence and/or loyalty. The advantage of this extension, like the previous ones, would be to offer a better match with the Polish case at hand, which is discussed in the next section.

In the following section, we take the main findings of the baseline model to data. To summarize, the most robust finding, even after accounting for the loyalty competency tradeoff, is that decreasing the competence of a typical former agent of the former authoritarian apparatus, makes selective purges the most attractive option **when the costs of verification are relatively low**. After introducing Poland, the setting of our empirical interpretation, we attempt to corroborate this finding with data from the Institute of National Remembrance on the operation of verification commissions in 1990.

4 Verification of security officers in Poland

In the late 1980s and early 1990s Poland underwent a transition from a communist authoritarian system to a free-market democracy. The transition was gradual and involved negotiations —the so-called "Roundtable Talks." Despite its name, the negotiation was bilateral and held between the old government and the from 6 February to 5 April 1989. The official Roundtable followed months of preparation, and included many secret rounds of bargaining over the terms of the negotiations. These latter talks took place in September 1988 in Magdalenka, a Warsaw suburb.

The topic of reforming the *Służba Bezpieczeństwa* (SB), as the communist security services were called, was even on the Magdalenka agenda. After all, it had been the SB who had been arresting and persecuting members of the opposition, some of whom had aspirations to become part of the Post-communist government. After a long history of repressing dissidents SB was perceived as a key pillar of authoritarianism and the future of the security services became a contentious topic in negotiations between Solidarity and the old government.

Reforming the SB presented a challenge due to its sheer size. In 1989 the Ministry of Internal Affairs employed 125,625 people and, among them, officers of the SB constituted 24,308 (Kozłowski 2019). Moreover, aside from an emphasis on removing officers who were tainted by persecuting opposition, any reform had to take into account challenges facing the new democratic state. One crucial feature of Poland's regime change was the dual nature of transition: (1) from authoritarianism to democracy and (2) from market socialism to capitalism. Poland had to deal not only with dismantling institutions of the authoritarianism but also with large scale property crimes and corruption brought about by the uncertainties and unequal access to resources left behind by the outgoing regimes and fragile market institutions. Democratic governments needed officers trained in fighting corruption and white collar crime and to some extent the former secret police agents were better positioned to do this than insufficiently trained outsiders. Hence, the new security system was bound to include at least some personnel trained under the authoritarian SB. The new democracy was in dire need of a screening mechanism.

To make the new services credible and implement a modicum of transitional justice (Kozłowski 2019) the first democratically sanctioned cabinet implemented a series of "verification" commissions. Although, all SB officers employed as of 31 July 1989 were fired, each had a chance to regain employment in the newly created Office for National Protection (*Urzad Ochrony Panstwa, UOP*) on two conditions:

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

2. they had to be recruited by UOP

This second condition was non-trivial, as while some 14000 officers participated in the verification process, there were only c.a. 5000 vacancies in UOP. Hence, verification was non synonymous with re-employment in the new secret service. The process of approval or removal was sanctioned by the Central Verification Commission alongside the Qualifying Commission, but initiated by one of 49 Regional Verification Commissions (RVC). Established in July 1990, RVCs were composed of a mix of staff from the old communist and new-democratic regime. They included a representative of UOP, the head of local police, universally trusted local activists as well as MPs, and senators from each *województwo* (as 49 the regions in Poland are called).

With little more than a few weeks to review hundreds of personnel file RVCs resources were thinly stretched. Unsurprisingly, they resorted to shortcuts: they relied on work assignments, flags of disciplinary issues, and eagerness to use violence to accelerate their evaluations. Moreover, the commissions adopted different criteria for evaluation. Some only approved officers who stood out by having achieved some socially desirable outcomes, while rejecting all remaining. Other commissions rejected only those who were proven guilty of transgressions, while retaining all other candidates. This contributed to the high variance in the proportion of officers verified negatively, as discussed in the coming sections.

The decisions of the RVCs were not final. Officers could appeal the outcome of their verification to the Central Verification Commission (CVC). The CVC either made their own independent decision relying on the personnel files or additional materials supplied by the candidate or requested that the relevant RVC to reconsider or provide further justification for its decision. Ultimately, out of the 14034 security officers that underwent verification a vast majority, 10439, were verified positively and could go on to apply to work in the new democratic security service.

4.1 Poland as a case study

Poland is an appealing case to study for several reasons. First, just as in the motivating case of the US, during its transition years it needed to deal with security forces that had loyalties and beliefs that could potentially destabilize the state. As described above, the new Polish democratic government had to decide how to retain competent moderate officers capable of staving off the threats of the transition while sifting through the disloyal and/or possibly incompetent"bad apples."

Moreover, Poland's 49 województwa offer considerable variance in the outcome of operations of the Regional Verification Commissions. The range of negatively verified officers runs from 7% to 80%, allowing us to see all three types of purges considered in the model within one country. A sub-national study like this allows us to keep constant other 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 in reminiscent of subnational analysis of, say, state legislatures in the 50 United States.

4.2 Data

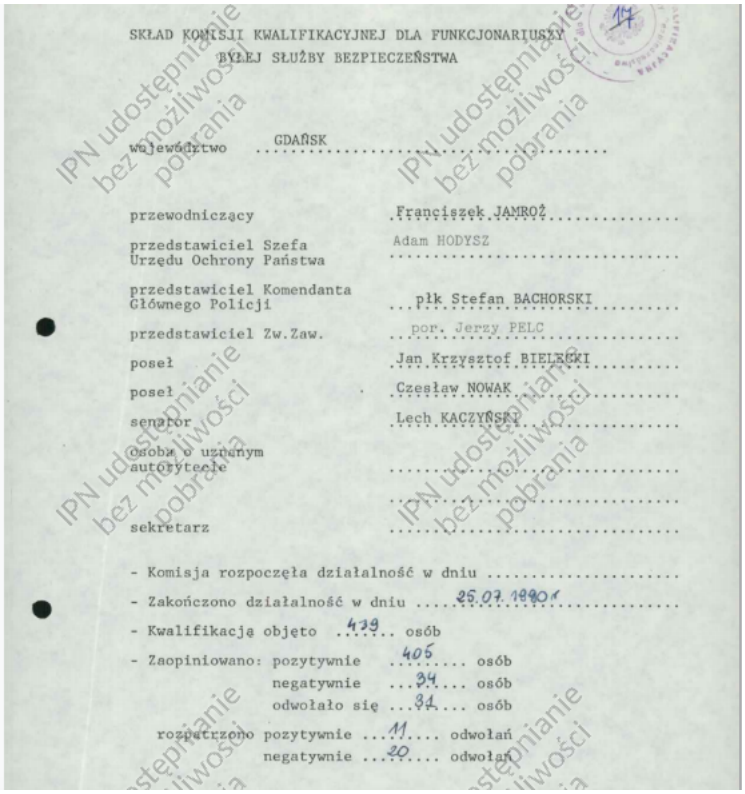
To understand the determinants of the type of purges, we have analyzed a number of historical and archival sources, some of which we compiled into original datasets. As discussed in the theory section, the key explanatory variables were the level of competence among officers in a given województwo, the extent of extremism among the officers, and the cost of conducting a selective—that is, associated with a careful reading of personnel files—verification commission.

The archival data from files of the Regional Verification Commissions (*Wojewódzkie Komisje Kwalifikacyjne*) allows us to extract from each of the 49 województwa[2] a number of measures concerning the severity of purges. These include the number of officers that

[2] We are currently waiting for files from the remaining 10 wojewodztwa to be delivered by the Polish Institute of National Remembrance (*Instytut Pamiéci Narodowej* or IPN).

have undergone verification, the number of those verified positively and negatively, as well as those that appealed the negative decision of the commission and the outcome of the appeal (Figure 1). The source of the archival data is the Polish Institute of National Remembrance and this is the first time to our knowledge that this data is used for quantitative analysis.

FIGURE 2: *An example of a Regional Verification Commissions summarizing the commission's decisions*



SKŁAD KOMISJI KWALIFIKACYJNEJ DLA FUNKCJONARIUSZY
BYŁEJ SŁUŻBY BEZPIECZEŃSTWA

województwo GDAŃSK

przewodniczący Franciszek JAMROŻ

przedstawiciel Szefa Urzędu Ochrony Państwa Adam HODYSZ

przedstawiciel Komendanta Głównego Policji płk 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 25.07.1990 r.
- Kwalifikacją objęto 439 osób
- Zaopiniowano: pozytywnie 405 osób
  negatywnie 34 osób
  odwołało się 31 osób
  rozpatrzono pozytywnie 11 odwołań
  negatywnie 20 odwołań

A cursory glance at the data from the verification commissions reveals a high level of variation in their level of severity. The proportion of those verified negatively ranges from 0.07 to 0.79 in the period preceding the appeals process (see Table 2).

FIGURE 3: *Outcome of verification*

histogram preappeal.png

Looking at the histogram (Figure 3) we can see that the voivodeships fell into three broad categories, as described in Table 1. We classify these wojewodztwa into three groups: (1)

TABLE 1: *Types of purges*

Purge type	Proportion of verified negatively	Frequency	Percent
No	<0.3	7	17.95
Selective	>0.3 and <0.7	27	69.23
Thorough	>0.7	5	12.82

having undergone no purge (NP), (2) a selective purge (SP) or (3) a thorough purge (TP) (see Table 1).

To study the effect of competence of the security services officer corps, we use data from an IPN publication covering the basic statistics on the type and number of officers at the regional level (Piotrowski 2003). The total number of officers prior to verification (1989) and the proportion of officers working in the security of economic resources (Department for the Protection of the Economy (PoE), *Wydział Ochrony Gospodarki*) in 1989 are the most relevant pieces of data. We use the latter as a proxy for competence, as the prestige of this department attracted highly skilled officers (Kozłowski 2019). These skills 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 that political in nature. According to our data, the proportion of PoE officers among all officers employed in the region varied between 0.13 and 0.53 (Table 2).

The level of extremism among officers in a given wojewodztwo is measured using the intensity of repression during the Polish Martial Law (1981-1983). Here, we focus on the number of people convicted of political crimes following the implementation of Martial Law

standardized by the number of Solidarity Trade Union members[3]. The standardization ensures the capture of the intensity of punishment conditional on the level of dissent. The data come from the current state of the IPN data on court cases against opposition members that took place during martial law (Instytut Pamieci Narodowej 2021). While the courts and the security services were nominally separate organizations, research by Popova (2012) and McCarthy (2015) suggests that in autocracies these actions of these institutions are more rigidly aligned. Hence, higher levels of repression by courts can be used as a proxy for extremism among the local security services.

Finally, to account for the cost of running the verification commission we use the density of security personnel relative to the population size. Województwa with smaller populations of had closely-knit networks and, implying a higher moral sense of justice towards those who carried out repression against dissidents. This mechanism has been highlighted by Krzysztof Kozłowski, the first chief of the Office of State Protection (*Urzad Ochrony Panstwa*), when he stated that "The regularity, more or less, is that the smaller the województwo, the more people know each other and the more informed the environment, that's where things happen that bear traces of reprisals [...] Whereas the larger the agglomeration, including the Ministry, there the names are without emotional overtones, it's more anonymous" (Kozłowski 2019, p.206). Hence, in less populous województwa, with denser networks of security officers the commission would pay a higher cost for engaging in a scrupulous examination of individual officer files. In those smaller, tightly knit województwa, the easy way out was to succumb to popular demand for revenge.

[3] The Solidarity Trade Union was the dominant anticommunist organization in the 1980's. Its exponential growth ultimately prompted the authorities to implement Martial Law.

TABLE 2: *Descriptive statistics*

Variable	Obs	Mean	Std. Dev.	Min	Max
Verified negatively	39	0.4766944	0.1980265	0.0671642	0.7913043
Protecting the economy	48	0.265018	0.0880872	0.1329114	0.5333334
Sentenced	47	38	50.78082	0	213
Solidarity members	49	196331.9	187110.9	33853	1122114
Security officers	49	148	94.12337	45	541
Population	49	761,814	596134.1	244300	3916400

5 Analysis

Given the above operationalization of the key variables, we can now check the model's propositions against the data.

As a reminder for the reader:

- μ - the proportion of competent agents, is measured as the proportion of officers who prior to 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 density of security officers in each wojewodztwo
- θ - 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

We interpret the policy shock ϵ as the uncertainty related to the regime transition and hence common for the whole country. This means that for the theoretical propositions ϵ can be held constant and so does not enter directly our empirical analysis and does not need to be operationalized.

According to the propositions of the model, we should (1) see more selective purges where the proportion of competent agents is small (2) see more thorough purges where *either* the proportion of competent agents is high and proportion of extremists low *or* proportion of competent agents is middling and the proportion of extremists is high. Increased cost of setting up the commission (3) unambiguously decreases the payoff from a selective purge.

In theory, given the tripartite nature of choice of the democratic policy maker, the empirical theory would ideally involve a Multinomial Logit (MNL) Model. However, given the low number of observations (N=39[4]), we start by analyzing the data using a simple comparison of means (Table 3).

TABLE 3: *Means of the purge type determinants*

Type of purge	competence	SD(competence)	repression	SD(repression)	cost	SD(cost)
No purge	0.27019	0.08567	0.12918	0.144663	0.78094	0.68613
Selective purge	0.24465	0.07057	0.20731	0.16626	0.48703	0.12417
Total purge	0.39581	0.10788	0.18703	0.19969	0.52942	0.09940

As postulated by the model, województwa that chose a selective purge are also the ones where the level of competence prior to transition was the lowest. The effect of the verification cost is also visible in the graph: the areas with lowest density of security services in population, that is those where the cost of setting up a commission was less costly, are also the ones that saw selective purges.

The relationship between competence and a thorough purge is non-linear, hence a simple comparison of means is not sufficient to verify the implications of the model. Table 4 presents the full data on competence[5] and extremism and the type of purge chosen in a given województwo.

We can see here that a thorough purge is generally more likely for higher levels of competence. Especially in the case of the most competent województwo, this is also accompanied with a low level of repression, as postulated by the model.

6 Conclusion

This paper has analyzed the dilemma facing new democracies in the aftermath of their transition to democracy. These new democratic governments must rebuild their ships while

[4] Eventually we hope to have all 49 wojewodztwa in the dataset

[5] Sorted by level of competence

as see, to use the words of Jon Elster et al. (1998). Specifically, we analyzed the policies of reforming and/or purging the security apparatus of the authoritarian regime. Without rigorous analysis, the answer to the question who should be fired seems obvious: the bad apples. But the meaning of "Bad" here is ambiguous. It can refer to lack of competence, but it can also mean ideologically distant from the new democratic government. In this latter case, retaining competent agents is actually detrimental to the new democracy.

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 will take the time to carefully verify their security apparatus through a process we have called a "selective purge." We show that as long as the costs of verification are not prohibitively high, selective purges are more likely when on average former security officers are less competent. This may seem puzzling at first, but becomes clear upon deeper reflection: if few agents are competent, it is simply be worth paying the extra verification cost to find out who exactly are the competent ones.

Our theoretical finding are corroborated with preliminary analysis using data from the operation of 49 regional verification commissions in Poland in 1990. We find preliminary support for our theory, lending further confidence in the findings of the baseline model.

Several extensions remain to be examined in a later draft of this paper. We conclude this version with a consideration of scope conditions.

It is worth considering to what extent can Poland is representative of screening mechanisms for former secret servicemen in the Post-Communist context. Poland shares the dual nature of the transition from authoritarianism and from a planned economy countries in the region, but what sets it apart from many is the very gradual process of the transition and the fact the the pace of transition itself was the subject of negotiation. It is very likely that the pace factored into the price of selective purges, making careful verification simply more affordable. One could however also make the opposite argument according to which gradual transitions allowed the outgoing regime to destroy personnel files of secret police officers,

making the ultimate work of verification commissions harder. Kozłowski (2019) cites numerous sources as evidence that such destruction of archives took place in the ten month long transition period, although most of the sources point to destroying files of informers, rather than officers themselves. Taken together, this suggests that selective, as opposed to thorough or no purges, could be more expensive in countries where the transition was more abrupt, perhaps even prohibitively expensive. Even despite these caveats, the underlying problem modeled here—namely, the combination of personnel shortages with acute uncertainty concerning competence and loyalty to the former authoritarian rulers among state employees—is universal. Even countries where the transition occurred without negotiation have to make personnel decisions, at least in the short term before new competent agents and bureaucrats can be trained. Universally, these new democracies need employees who can "uncover the state of the world" for them. Hence, the lessons from the theoretical model are applicable beyond this specific case.

References

Callander, S. et al. (2008), 'A theory of policy expertise', *Quarterly Journal of Political Science* **3**(2), 123–140.

Debs, A. (2016), 'Living by the sword and dying by the sword? leadership transitions in and out of dictatorships', *International Studies Quarterly* **60**(1), 73–84.

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.

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 Pamieci 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, M. M. (2019), 'Generalized backward induction: Justification for a folk algorithm', *Games* **10**(3), 34.

Kaminski, M. and Nalepa, M. (2014), 'A model of strategic preemption: why do post-communists hurt themselves?', *Decyzje* (21), 31–65.

Kozłowski, T. (2019), *Koniec imperium MSW: transformacja organów bezpieczeństwa państwa 1989-1990*, Instytut Pamieci Narodowej–Komisja Scigania Zbrodni przeciwko Narodowi ....

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.

McCarthy, L. A. (2015), *Trafficking Justice: How Russian Police Enforce New Laws, from Crime to Courtroom*, Cornell University Press.

Nalepa, M. (2021), *After Authoritarianism: Transitional Justice and Democratic Stability*, Cambridge University Press.

Nozick, R. (1978), *Anarchy state and utopia*, Blackwell, London, England.

Olimpieva, E. (2021), *Putin's Prosecutors: How Law enforcement Helps Build Authoritarian States*, Unpublished Manuscript.

Paine, J. (2022), 'Reframing the guardianship dilemma: How the military's dual disloyalty options imperil dictators', *American Political Science Review* **Forthcoming**.

Piotrowski, P. (2003), 'Struktury Służby Bezpieczeństwa MSW 1975-1990', *Pamieć i Sprawiedliwość.* **3**(1), 51–108.

Popova, M. (2012), *Politicized justice in emerging democracies: a study of courts in Russia and Ukraine*, Cambridge University Press.

Woldense, J. (2018), 'The ruler's game of musical chairs: Shuffling during the reign of ethiopia's last emperor', *Social Networks* **52**, 154–166.

Zakharov, A. V. (2016), 'The loyalty-competence trade-off in dictatorships and outside options for subordinates', *The Journal of Politics* **78**(2), 457–466.

TABLE 4: *Levels of key variables and purge types*

wojewodztwo	purge type	competence	repression
Nowosadeckie	SP	0.141791	0.0184887
Lubelskie	SP	0.1531532	0.1360691
Piotrkowskie	SP	0.1538462	0.0444444
Wrocławskie	SP	0.1609195	0.5239562
Kieleckie	SP	0.1717172	0.25
Ciechanowskie	NP	0.1927711	0.0119464
Przemyskie	NP	0.1935484	0.027027
Radomskie	NP	0.1986755	0.0074074
Toruńskie	SP	0.2035398	0.3856172
Leszczyńskie	SP	0.2040816	0.0718236
Olsztyńskie	SP	0.2042253	0.1863729
Rzeszowskie	SP	0.2134146	0.2541667
Kaliskie	SP	0.2234043	0.0866667
Sieradzkie	SP	0.2244898	0.0294118
Koszalińskie	SP	0.2246377	0.3083333
Ostrołeckie	SP	0.225	0.1510058
Pilskie	SP	0.2293578	0.135129
Krośnieńskie	SP	0.2298851	0.0390165
Białostockie	SP	0.2318841	0.2318655
Siedleckie	SP	0.2417582	0.024724
Łódzkie	SP	0.2465753	0.4643481
Warszawskie	NP	0.2472406	0.2637277
Jeleniogórskie	SP	0.2542373	0.0316943
Czestochowskie	TP	0.2702703	0.150107
Gorzowskie	SP	0.272	0.3149606
Krakowskie	SP	0.2783019	0.499984
Legnickie	SP	0.2941177	0.1681024
Słupskie	NP	0.3023256	0.2
Elblaskie	SP	0.3240741	0.119403
Bielskie	TP	0.3309859	0.1111111
Gdańskie	NP	0.3461539	0.3668648
Bialskopodlaskie	SP	0.3484848	.
Bydgoskie	SP	0.3630573	0.292517
Płockie	SP	0.3645833	0.5452293
Chełmskie	TP	0.3661972	0.5317106
Łomżyńskie	NP	0.410628	0.0273115
Konińskie	SP	0.4230769	0.0767731
Suwalskie	TP	0.4782609	0.1290323
Skierniewickie	TP	0.5333334	0.0132073

NP signifies no purge, SP a selective purge, and TP a total purge