

Secret Authoritarian Legacies, Transitional Justice, and the Quality of Representation

Milena Ang
The University of Chicago

Monika Nalepa
The University of Chicago

Abstract

(150 words)

We formalize the classical “forward-looking” argument in support of personnel transitional justice (lustration). This argument states that by revealing former secret collaborators of the authoritarian regime among politicians, lustration prevents former authoritarian elites from extorting policy concessions from elected politicians. Absent lustration, former agents of the secret police could blackmail former collaborators-turned politicians and threaten to reveal skeletons in their closets if politicians refused to implement policies their former agents desire. Hence, implementing lustration laws should allow all politicians to become responsive to their constituencies. We show that whether lustration indeed prevents such distortions in representation depends on its severity, the extent to which dissidents-turned politicians face consequences if their skeletons “come out of the closet,” and media freedom. Surprisingly, the ideological distance between the successor autocrats and former dissidents is not relevant. We test this theory with an original transitional justice dataset covering 42 countries that transitioned to democracy since 1945.

Acknowledgments: The authors are grateful to Jaimie Bleck, Anna Grzymala-Busse, Karrie Koessel, Susan Stokes,¹ and Susanne Wengle for comments and to John Patty and Scott Tyson for advice on the formal model. The following research assistants contributed to compiling the original data: Evgenia Olimpeva, Tara Kelley, Alexa Rosario, and Yonatan Litwin. All mistakes are the authors’ responsibility.

1 Introduction

In February 2016, the widow of Czesław Kiszczak, former Chief of the Communist Secret Police in Poland, discovered a thick secret police file containing evidence that Lech Wałęsa, former Nobel Peace Prize Laureate and anti-communist dissident leader, had collaborated with the Communist Secret Police between 1970 and 1976. The revelation raised many concerns about the quality of Wałęsa's presidency, which ended before Poland implemented its lustration law. A lustration program would have vetted Wałęsa for connections to the communist secret police. Had it been in effect when Wałęsa ran or held office, his political career could have ended if he erroneously maintained his innocence. Although Wałęsa's collaboration preceded his career as dissident and trade union organizer, Kiszczak might have pressured him to avoid implementing certain policies by threatening to release the compromising file. This led many people to question the quality of his presidency and the extent to which he represented interests of the electorate rather than those of the former secret police. More generally, it left scholars wondering about the prevalence of such acts of blackmail. How often did former authoritarian elites pressure their former spies into following their policy preferences instead of representing voters? To the extent that we consider representing voters' wishes a marker of high quality representative democracy, the revelation left many wondering just how representative Polish democracy was of its electorate.

Political scientists do not dispute that the power of former authoritarian elites extends beyond the life span of an authoritarian regime. This can happen for two reasons. First, autocrats are better positioned to capture state resources prior to the democratic transition, which they can then use in a clientelistic fashion to stay in power (Brun & Diamond 2014). The outgoing autocrats access to resources can be cut off if they or their successors are voted out of power following the transition to democracy, but as various cases around the world demonstrate (Kitschelt 1999), this removal is frequently just temporary. Scholars have attributed the revival of successor authoritarian parties to the organizational advantage they hold over parties that are new to the party system. This organizational advantage allows

them to make better use of state resources when they eventually do find themselves in government (Grzymala-Busse 2002). Second, autocrats may retain power after the transition, because the transition itself is endogenous to their strength. In their recent article, Slater and Wong explain the advantage of former authoritarian elites following democratic transitions by arguing that they concede their power only in moments when they can expect to thrive under democratic conditions at least as well as they did under autocracy (Slater & Wong 2013).

Given these two sources of power the former autocrats have— one is exogenous and the other endogenous to the transition itself— the only institutional undercutting this unfair advantage is Transitional Justice (TJ). Scholars of TJ have argued that these laws can undercut the privileged position of members of the former autocrats, their parties, enforcement apparatus as well as its collaborators (Stan & others 2009, David 2011).

Transitional Justice comprises of the “formal and informal procedures implemented by a group or institution of accepted legitimacy around the time of transition out of an oppressive or violent social order, for rendering justice to perpetrators, and their collaborators, as well as victims” (Kaminski, Nalepa & O’Neill 2006). The four pillars of transitional justice are truth commissions trials, victim compensation, and lustration. If the normative objective is barring former authoritarian elites from policy influence, then the transitional justice mechanism of paramount importance is lustration, which vets candidates for public office for ties to the former authoritarian secret police. It does this through opening archives of the secret police of the former authoritarian regime to uncover who among persons running for public and political office had worked for the secret police prior to the democratic transition. The proven collaborators are then either explicitly banned from holding office or information about their collaboration is released to the voters, who then can decide whether to still cast their votes on thus compromised politicians.

There are many retrospective or backward-looking arguments in support of lustration, ranging from victims of spying having the right to know who informed on their activities to the authorities (Stan 2012) to the need to prevent former spies and their leading officers from

playing key roles in public service (Nedelsky 2013). Yet some scholars dismiss transitional justice for this retroactive character and argue that “living well is the best form of revenge” (Halmai, Scheppelle & McAdams 1997). Among the advocates of this so-called “Spanish Model of Transitional Justice” are Elster (1998) and Holmes (1994), who used Spain’s most benign way of dealing with former authoritarian collaborators (sealing off the archives of Franco’s secret police) to build their case that “doing nothing” is the best approach for new democracies to deal with past authoritarian regimes (Elster 2004). Extensive scholarship on personnel transitional justice corroborates this endorsement by showing misuses of lustration, de-communization, and de-ba’athification. To take one example, Iraq’s policy aimed at purging new democratic institutions of former Ba’athists to promote societal reconciliation. De-ba’athification prevented 185 members of Saddam Hussein’s party, mostly Sunnis, from running for the legislature in 2003, but instead of reconciliation it ignited ethnic tensions (David 2006). France’s policy of *épuration* banned former Vichy collaborators from holding office following WWII had very similar effects (Kritz 1995, Elster 2004).

However, although following the Spanish model and “doing nothing” need not produce immediate negative consequences, it may strengthen the power of authoritarian networks. If democracy survives, embarrassing information collected by the former authoritarian secret police for the benefit of authoritarian elites may turn elected politicians into clients of agents who threaten to reveal their “skeletons in the closet” (Nalepa 2010). Forgiving and forgetting may sabotage the capacity for elected politicians to represent voters, a phenomenon that is hard to pick up on for studies focusing on the immediate aftermath of transition. Thus, even if one agrees with the arguments of Elster and Holmes questioning the validity of the backward looking arguments for lustration, it may still be the case that lustration laws can, when appropriately designed and implemented, undercut the growth of authoritarian networks by exposing the files and secrets of the former authoritarian police. According to this argument, revealing evidence of human rights violations and collaboration with members of the ancien regime prevents former authoritarian elites from influencing policy in new

democratic polities. Where lustration is lacking, former authoritarian elites can pressure politicians into policy concessions by threatening to reveal compromising information that could jeopardize their careers.

Our paper seeks to examine in more detail this forward-looking argument for transitional justice. We develop a theoretical model reconstructing the blackmail mechanism that lustration is supposed to undercut. By making public the secret information that former autocrats could use to influence policy-making in the new democracy lustration should improve programmatic representation in new democracies. We use a game theoretic model of incomplete information to reconstruct the mechanism through which lustration laws prevent political participation of persons whose dependence on the former security apparatus interferes with their ability to faithfully represent policy preferences of the electorate.

Our first contribution is to uncover circumstances under which lustration will indeed improve the ability of politicians to represent preferences of their constituents as opposed to their secret agents. The crux of our theoretical contribution is identifying conditions under which agents extract policy concessions from politicians even when lustration is in force. We identify (pooling and semi-pooling) equilibria under which lustration is compatible with bluffing and consequently, departures from democratic representation. We also present an identity result, which shows that irrespective of the equilibrium being played, the formula describing misrepresentation is the same across all types of equilibria.

Our second contribution to the literature on transitional justice is empirical. In order to test the implications of our theory we have created a novel dataset comprising of a time series cross-section of personnel transitional justice events in 42 post-authoritarian states. Despite a growing number of empirical studies examining the broad impact of lustration and transitional justice on democratic stability and peace (Olsen 2010, Van der Merwe, Baxter & Chapman 2009, Thoms, Ron & Paris 2008), there is a knowledge gap pertaining to the impact of personnel transitional justice policies, such as lustration, on the long term quality of democratic representation. The first generation of empirically oriented transitional justice research

was concentrated on explaining patterns and trajectories of implementing specific transitional justice measures (Stan & others 2009, Lundy & McGovern 2008, Mallinder 2008, Pettai & Pettai 2014). More recently, to understand the impact of transitional justice on the quality of democracy, scholars have begun to investigate transitional justice as an independent variable. Noteworthy because of their coverage of lustration policies are contributions by Horne (n.d.) and Olsen (2010).¹ Yet while Olsen (2010) aggregates all lustration activity into a dummy variable summarizing the presence or absence of lustration following the transition Horne (n.d.), provides a rich description of lustration trajectories for as many as 12 countries, but only in Post Communist Europe.²

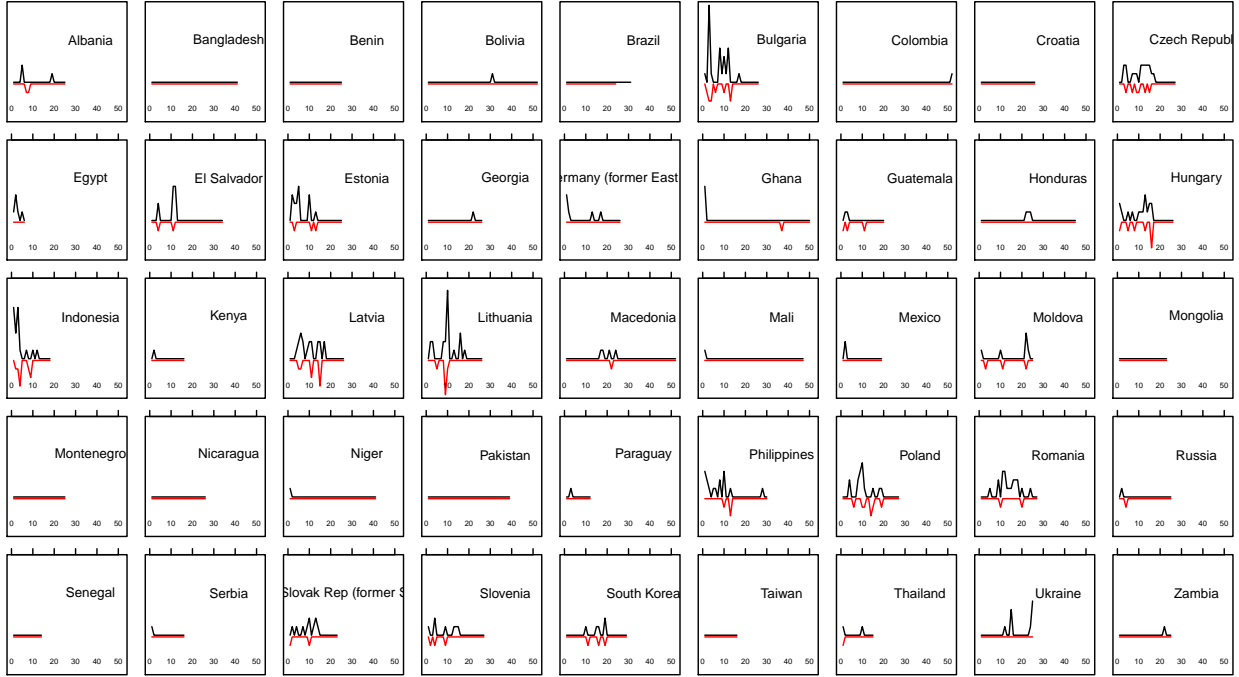
Figure 1 summarizes the data we use to construct our dependent variable. It is a time series of positive and negative transitional justice events per year since the democratic transition in forty two countries that transitioned from party-based authoritarianism to democracy since WW 2.³ It shows first that limiting data collection on lustration to Post-Communist Europe is unwarranted. Searching for personnel transitional justice using keyword searches broader than the term “lustration” reveals a much richer scope of personnel transitional justice.

¹Horne examines the relationship between the nature of transitional justice mechanisms and trust, suggesting that revealing the notoriety with which citizens were spying on one another may decrease interpersonal trust in countries with a large network of covert collaborators. At the same time, her research indicates that a wide and compulsory lustration procedure may result in substantial bureaucratic turnover, thereby increasing political trust in governmental institutions. Horne’s research was among the first to disaggregate trust in this manner; much of the previous literature assumed that transitional justice was essential for trust-building.

²Indeed, most contributions to the lustration literature suggests that this transitional justice policy is almost exclusively restricted to Post Communist Europe (Mayer-Rieckh & De Greiff 2007, Ellis 1996, Closa Montero 2010, Letki 2002, Volčič & Simić 2013, Stan 2013, Stan & Nedelsky 2015). This is however an artifact of restricting the search for personnel transitional justice to the term “lustration” and its derivatives, when in fact purging the state apparatus of members of the former authoritarian regime and their collaborators is called by different names in different parts of the world, from “vetting” and “purging” to “house-cleaning” as well as specific terms, such as “de-nazification,” “de-communization,” or “de-ba’athification.”

³A positive transitional justice event (in black) signifies the submission of a lustration proposal to the floor of the legislature, the passage of such legislation, the upholding of such legislation as constitutional by a supreme court, or the overturning of a presidential veto against such legislation. A negative transitional justice event (in red) signifies the voting down or vetoing of lustration provisions by the constitutional court. Similarly, the expansion of persons targeted by lustration or broadening the set of lustrable offenses constitutes a positive transitional justice event, whereas attempts to narrow the set of targets or lustrable offenses were coded as negative transitional justice events.

Figure 1: Summary of Negative and Positive Events in Personnel Transitional Justice



The limitations of reducing a sometimes very complex process of shepherding a transitional justice proposal through the legislative process to a dummy variable is underscored by the fact that while in some countries the bulk of lustration activity occurs in the immediate aftermath of transition (as in Kenya and Ghana), in others, the process is far more protracted (as in most Eastern European countries). Furthermore, positive transitional justice events seem to prevail in certain countries (as in Ukraine and for a long time in the Philippines), whereas in others they are accompanied by reversals (as in Indonesia).

The next section introduces our formal model and presents the identity result and the empirical implications of the model. Section 3 presents the empirical strategy for evaluating two hypotheses about programmatic representation. Section 4 concludes. All proofs are relegated to the formal appendix.

2 The Formal Model

Formal modeling has scarcely been used to address questions involving transitional justice⁴ and to the best of our knowledge has not been used in predicting the effect of personnel transitional justice mechanisms on the long term quality of democracy.

Can lustration limit the influence of former authoritarian elites? We approach this question by developing a formal theoretical model that reconstructs the mechanism through which former secret police officers influence politicians' policy-making decisions. Solving this model allows us to examine to what extent transitional justice indeed reduces the dependence of politicians on their former leading officers. These former leading officers are referred to here as "agents."

Our model indicates that, contrary to suggestions from the literature: (1) the degree of ideological polarization between the secret police agent and the politician is inconsequential for programmatic representation; (2) programmatic representation increases with the severity of lustration and (3) it is easier to achieve as the price of having skeletons in the closet revealed decreases.

There is also an interesting implication for the role of free media on lustration blackmail: as media becomes more free and capable of scrutinizing false information about skeletons in the closet (or bluffing), lustration blackmail occurs only in instances where evidence does in fact exist. In other words, in a world of free media, transitional justice works as intended: blackmail only transpires when evidence of collaboration is left in the hands of former members of the authoritarian regime. The intuitive logic driving this finding is as follows: former collaborators-turned-politicians cannot know with certainty if files against them survived the transition, but former agents can bluff politicians into thinking such evidence exists if their bluff cannot be easily challenged. Free media, however, can expose such a bluff more easily. Hence, without free media, former agents have the ability to extort policy concessions from politicians even when they actually do not have skeletons to reveal.

⁴Exceptions include Ritter & Wolford (2012), Nalepa (2010) and Casper & Tyson (2014)

In equilibria that allow for bluffing (pooling and semi-pooling equilibria) blackmail is effective despite lustration. These are contrasted with separating equilibria, where blackmail is only effective when evidence was left behind by the lustration procedure.

A possible criticism of how well our model matches the empirical interpretation of blackmail with secret police files is that surely the collaborator himself knows whether evidence of his or her collaboration exists or not. If that is the case, then “bluffing with secret police files” is limited to politicians who collaborated with the secret police and those who are innocent of such collaboration cannot be affected. We point out that in light of falsified evidence of collaboration that secret police In response, we point to the extensive literature indicating that secret police agents routinely falsified evidence of collaboration in order to improve the appearance of their performance and to pocket compensation that was intended to be given to collaborators in exchange for information (Horne 2009, David 2003). Not infrequently were informal conversations recorded as “reports.” The victim of such falsified evidence could never be certain therefore is evidence of the “collaboration” exists or not (Nalepa 2008). Thus, because of such instances of unintentional collaboration, we assume that theoretically, bluffing was always possible.

Our model features a former collaborator who is elected to public office following a transition to democracy and his agent who threatens to reveal embarrassing information against the former collaborator, unless the latter implements policies favored by the agent. The model assumes that, absent the pressure of blackmail, politicians would programmatically carry out the policies their voters’ desire. Democratic representation is undermined when the politician is vulnerable to extortion by the agent and implements the policy preferred by his former secret police officer, instead of the policy preferred by voters.

The key tension in our model comes from the fact that while the agent knows whether or not he possesses the evidence he threatens to release, the former collaborator does not. Notice, that the more stringent the lustration law the less likely it is that such incriminating evidence is in possession of the former agent. Therefore, we model the severity of lustration

as the probability that evidence of the politician’s collaboration remains in the possession of the agent. Returning briefly to our example from the introduction, if harsh lustration had been implemented in Poland before Wałęsa served as President, the secret police chief would not have been able to use the files against Wałęsa.

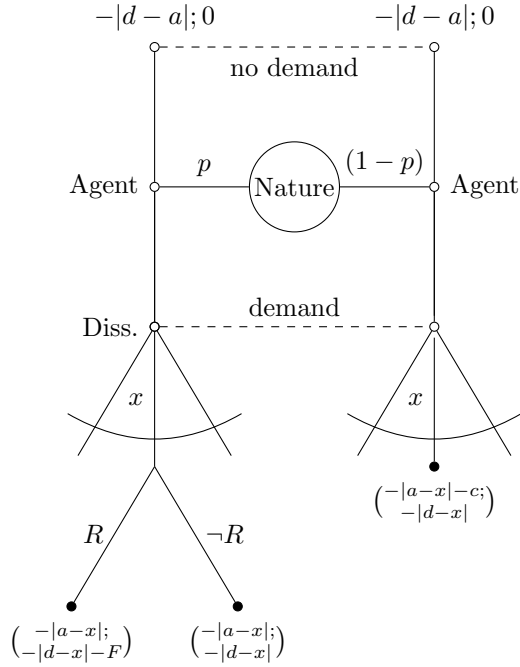
2.1 Sequence of play

The game starts with a move of Nature which with probability $p \in (0, 1)$ determines that evidence against the collaborator is in the possession of the agent and with probability $1 - p$ that it is not. The politician, who we will call the *Dissident* (D), knows the value of p , but not its specific realization.⁵ In the second stage of the game, the *Agent* (A), with an ideal point $a = 0$, decides whether or not to make a policy demand towards the *Dissident*. For simplicity, we represent his decision by one of two exogenously given actions {Make Demand, Not Make Demand}. In the third stage, the *Dissident* with ideal point $d \in (0, 1)$, who observes whether or not a demand on him has been placed, but not whether evidence against him exists (the realization of p) decides whether to make a policy concession in response to the Agent’s demand and also how big of a concession to make. This is modeled as proposing a policy $x \in [a, d]$. If evidence against the Dissident does not exist, the game ends here. But if evidence does exist, the *Agent* decides whether or not to reveal the evidence against the Dissident. The idea here is that if the concession is deemed insufficient by the Agent, he can exercise his threat and indeed reveal the compromising evidence against the Dissident.

The payoffs are a function of the Euclidean distance between the players’ respective ideal points and implemented policy. We also assume that the *Agent* pays a cost $c > 0$ for making a demand when he does not have evidence in his possession. This makes sense, as trying to coax a politician into believing that the agent has evidence against him may require obtaining and disclosing some plausible information about the *Dissident*’s past which could

⁵We refer to the ex-collaborator turned politician as “dissident” to avoid confusing his ideal point (d) with the probability that secret files are in the possession of the agent. This nomenclature is also justified by the fact that in places where collaboration with the secret police was of clandestine nature, as was the case in PostCommunist Europe, the revelations are most damaging to former dissidents (Nalepa 2010)

Figure 2: Lustration Blackmail Game



be verified by the media, if the media has sufficient independence.⁶

In addition, if evidence against him is revealed, the *Dissident* pays the cost of being fired, F . The utilities, along with the entire game tree are presented in Figure 2 above.

2.2 Strategies and beliefs

A strategy for the agent is a triple $s_A = (z^E, r, z^{\sim E})$, where $z^E \in \{0, 1\}$ denotes A 's action when he is in possession of evidence against D and $z^{\sim E} \in \{0, 1\}$ denotes A 's action when he is not in possession of such evidence, with action 1 denoting “make demand” and action 0 denoting “do not make demand.” Suppose that $r \in \{Reveal, \sim Reveal\}$ denotes the action taken by A following D 's counteroffer x . This part of A 's strategy can best be represented as a rejection region $R = \{x : x \in R \rightarrow A \text{ reveals evidence}\}$. A strategy for the dissident is

⁶We note here, following (Nalepa 2010) that many dissidents were unconscious collaborators in that they met with undercover agents in their daily lives and provided them with important information about their dissident activity unknowingly. Sometimes this seemingly innocuous information would allow the secret police to advance their repression of the anti-authoritarian opposition and from the point of view of a politician revelation of it could be just as damaging.

a proposal $x \in [a, d]$ that D makes in the event that A makes a policy demand. Note that D can simply ignore the demand if he chooses $x = d$.⁷

This structure resembles that of a signaling game, in which A is the sender of the message, while D is the receiver. In line with these types of games, the dissident may have a chance to update his prior belief about the type of agent he is facing—whether it is an agent equipped in evidence of collaboration or not—after observing the agent’s action. Depending on the relative magnitudes of d, c and F , this might be a signaling game or a cheap talk game. We will assume, not without loss of generality, but in order to analyze the most interesting cases of this general setup, that $0 = a < d$. The solution concept we use is that of Perfect Bayesian Equilibrium (PBE). This means that in addition to strategies being best responses to each other, we also have to specify the *Dissident’s* beliefs and those beliefs must be consistent with what is observed in equilibrium play.

2.3 Results and Discussion

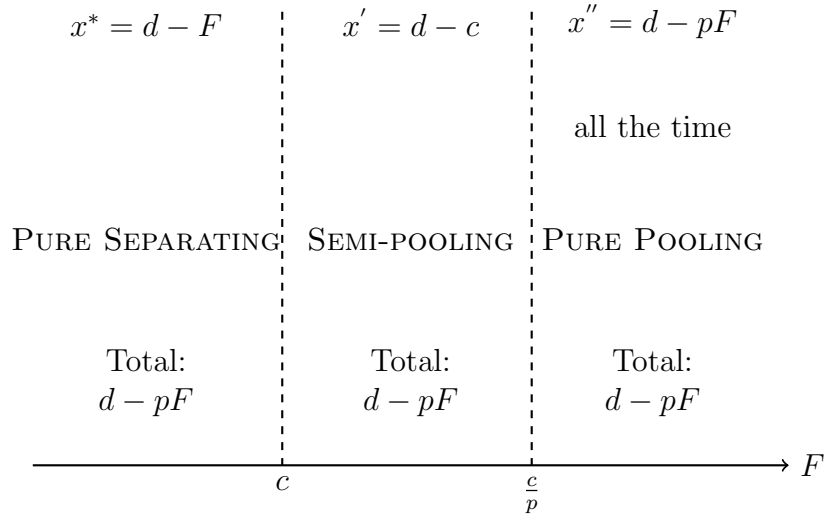
The three types of equilibria are presented in proposition 2.1 as well as in Figure 3 as a function of F , the dissident’s cost of being fired. The set of conditions for the pooling equilibria is distinct from the set of conditions defining the signalling equilibria, with the critical element being the magnitude of F relative to the cost of bluffing, c and the probability that evidence against the dissident exists, p .

Proposition 2.1 1. Let $x^* \equiv d - F$. For $F < c$ and the posterior beliefs of D be described by $Pr(E|0) = 0, Pr(\sim E|0) = 1, Pr(E|1) = 1, Pr(\sim E|1) = 0$, $(0, R, 1; x^*)$ is a (pure separating) PBE.

2. Let $x'' \equiv d - pF$. For $F \geq \frac{c}{p}$ and the posterior beliefs of D be described by $Pr(E|1) = p, Pr(\sim E|1) = 1 - p$ $(1, R, 1; x'')$ is a (pure pooling) PBE

⁷Note that since D does not observe whether evidence exists or not, but only observes whether a demand was placed (0) or not (1), we only need a single action do describe his strategy.

Figure 3: Equilibria



3. Let $x' \equiv d - c$ and $\lambda^* \equiv \frac{p(F-c)}{(1-p)c}$. For $c \leq F < \frac{c}{p}$ and the posterior beliefs of D be described by $Pr(E|0) = 0, Pr(\sim E|0) = \lambda^*, Pr(E|1) = 1, Pr(\sim E|1) = 1 - \lambda^*$, $(1, R\lambda; x')$ is a (semi-pooling) PBE.

The derivation of pure separating, semi-pooling, and pooling equilibria is provided in the Formal Appendix. Below we also present an identity result characterizing the PBE outcome, followed by a discussion of this result.

Recall, that our model assumes that absent pressure from the agent, the dissident would carry out the policies the voters desire, which correspond to his ideal point, the programmatic platform the dissident's party ran on in the elections. Ultimately, what we are most interested in across all equilibria is the level of programmatic representation: the extent to which the dissident can withstand pressures from the agent to abandon his ideal point, d . In addition to the values of x^*, x'' , and x' , which correspond to the counteroffer proposed by the dissident (in the the separating, pooling and semi-pooling equilibria respectively), we present the average levels of programmatic representation. We interpret programmatic representation as resistance to the agent's blackmail. Conversely, misrepresentation is how we interpret the departures from d , the Dissident's ideal point.

In order to derive the expected level of programmatic representation we weigh the PBE outcome in each equilibrium by the frequency of its occurrence. In the pooling equilibrium, x'' is played all of the time. Hence, the total does not need to be weighed. However, in the pure and semi-pooling equilibria, to arrive at the total expected value of the counterproposal, we need to apply weights corresponding to the probability with which each outcome in equilibrium will be reached.

In the case of the pure separating equilibrium, it is simply $p(d - F) + (1 - p)d$, as the agent only proposes $x^* = d - F$, when evidence exist, which is p of the time. The remaining $1 - p$ of the time, he reverts to d . In the case of the pure pooling equilibrium, the programmatic representation is $1 * (d - pF)$, as the agent always places a demand and the dissident always responds with $x'' = d - pF$. In the case of the semi-pooling equilibrium, the calculation of programmatic representation is somewhat more complex, because the agent places a demand when evidence exists, p , and λ of the time when it does not exist. Hence the total frequency of placing a demand is given by $p + (1 - p)\frac{p(F-c)}{(1-p)c}$. The dissident responds to this demand with x' , bringing the total of programmatic representation to $(d - c)(p + (1 - p)\frac{p(F-c)}{(1-p)c}) + (1 - p)(1 - \frac{p(F-c)}{(1-p)c})d$ which as in the previous two cases reduces to $d - pF$. These totals are presented as a function of d and p in Figure 3 and described below in Proposition 2.2, which follows directly from our derivation above.

Proposition 2.2 *The PBE outcome, interpreted as the expected level of programmatic misrepresentation, is given by the same formula across all three equilibria. $d - pF$.*

A few implications follow from proposition 2.2. First, departures from perfect representation are constant for all values of d . This means that the effectiveness of blackmail does not depend on how far apart the ideal points of the agent and dissident are located. This is somewhat surprising as intuitively, we would expect the ideological proximity of ideal points to matter. ⁸ Second, it is clear that the quality of representation is decreasing in p . This is

⁸As a robustness check, we solve the same model but using quadratic as a opposed to linear spacial preferences and for two of the equilibria, the effect of d was negative

very intuitive as the probability that evidence does not exist induces the dissident to make a counteroffer closer to his ideal point. Third, programmatic representation is highest under the pure separating equilibrium, followed by the semi-pooling equilibrium, followed by the pure pooling equilibrium, indicating that it decreases with the cost of firing.

Finally, although the degree of programmatic representation is the same under the separating equilibrium as under the semi-pooling equilibrium, and especially under the pooling equilibrium, transitional justice does not prevent blackmail according to the logic its proponents would like it to. In other words, programmatic representation is achieved for the “wrong reasons.” This can be explained as follows. Starting with the pure separating equilibrium, bluffing never occurs and the amount of misrepresentation tracks the amount of evidence left in the agents’ possession (so it is directly responsive to the severity of lustration). Thus, under the separating equilibrium lustration works exactly as advocated by the proponents of the forward-looking argument for lustration. This equilibrium becomes easier to achieve as the costs of bluffing increase, because for higher values of c , it is easier for F to be low enough to satisfy the $c < F$ constraint.

In the semi-pooling equilibrium $(0, R, \lambda; x')$, the Agent always makes a demand whenever the evidence exists and also when evidence does not exist with probability λ^* . The key feature of the semi-pooling equilibrium is that when it is played, the agent extracts policy concessions, undermining representation, even when evidence does not exist with probability λ^* . Misrepresentation does not exactly track transitional justice, but at least in $1 - \lambda^*$ of the cases when evidence has been swept by the transitional justice process, bluffing does not take place. This equilibrium occurs when $c < F < cp$.

In the pooling equilibrium, bluffing occurs every time evidence does not exist regardless of how much evidence transitional justice has left in the hands of agents. Either the separating or the semi-pooling equilibrium is better from the normative point of view than the pooling equilibrium, where the frequency of blackmail is completely orthogonal to the actual evidence of collaboration. Under this equilibrium, the extent to which lustration has success-

fully eliminated evidence of collaboration has no bearing on the effectiveness of blackmail. Blackmailing or bluffing occurs frequently and lustration is completely dysfunctional. This last equilibrium occurs when the possibility of having skeletons in the closet revealed by the former secret police is very damaging to the politician's career ($F > \frac{c}{p}$).

Consequently, although it is clear that raising c has no effect on total programmatic representation, a higher c makes the $F < c$ and $F < \frac{c}{p}$ constraints harder to satisfy. Thus, it makes the separating equilibrium more likely vis. a vis. the semi-pooling equilibrium or the semi-pooling equilibrium more likely vis. a vis. the pure pooling equilibrium. As a result, raising c makes transitional justice work more like the advocates of the forward-looking argument for lustration want it to work.

3 Empirical Analysis

We test the empirical implications of the model using an original dataset of partisan programmaticness in countries that transitioned from party-based authoritarian rule. Since our argument is that effective blackmail results in departures from policies that represent the electorate, the outcome of interest is the degree to which a party implements policies that correspond to the platform that voters elected it for. Hence, effective blackmail is observable as a decrease in a party's programmaticness, whereas parties that are not subjected to blackmail by former authoritarian elites should, on average, be more programmatic.

Our dependent variable builds on the measure of programmaticness constructed by Herbert Kitschelt and his co-authors (Kitschelt, Freeze, Kolev & Wang 2009, Kitschelt & Freeze 2010). Their original measure was constructed from a large-scale survey of experts in 88 different countries answering questions about the cohesion, salience, and polarization of 506 parties. Programmaticness was conceptualized as building on these three aspects and an index was constructed based on responses to questions tapping into each of these components. Given the proximity between polarization and the operationalization of d as the

ideological distance between the successor authoritarian party and each non-successor party, we chose to create a new measure that leaves out questions pertaining to polarization. The result is a score between 0 and 1 that comprises just the cohesion and salience dimensions of the original variable. Programmaticness is measured at the party level, as in Kitschelt et. al..

The first implication of our formal model is that the probability with which evidence against the dissident exists (parameter p in the model) decreases the degree of programmatic representation. To operationalize p , we take advantage of lustration policies. As was discussed, lustration unearths existing evidence of former collaborators of the authoritarian regime, in an effort to prevent them from entering into politics. Thus, we argue that the more severe lustration, the lower the value of p ought to be. Put differently, the extent to which lustration within a country has been successful corresponds to the probability that evidence of collaboration remains in the hands of some former agent. Hence measuring the severity of lustration is the most natural operationalization of parameter $1 - p$.

Quantitative lustration data rely on indicator variables to measure lustration per country. However, this overlooks relevant variation in the severity of lustration policies implemented. In order to solve this, we present a continuous measure, obtained from original data. Using the Keesings Record of World Events, Lexis Nexis Academic Universe, and numerous secondary sources, we constructed a dataset of chronologies of personnel transitional justice⁹ events for former party-based authoritarian countries that transitioned to democracy since World War II. Each event was coded as either positive, or negative. Examples of positive transitional events are submissions of a lustration proposal to the floor of the legislature, the passage of such legislation, the upholding of such legislation as constitutional by a constitutional court, or the overturning of a presidential veto against such legislation. An example of a negative transitional justice event, in contrast, is the voting down, vetoing or strik-

⁹The literature defines lustration using three parameters: All persons in set X are screened for committing action Y in the past, and if this procedure identifies a person in X responsible for engaging in action Y, he or she faces a sanction Z (Nalepa 2010).

ing down by the constitutional court lustration provisions. Similarly, expanding the set of persons targeted by lustration or broadening the set of lustrable offenses to include more past or present positions constitutes a positive transitional justice event, whereas attempts to narrow the set of targets or lustrable offenses are coded as negative transitional justice events

We aggregated the number of positive and negative transitional justice events to create a panel, with countries as the cross section and time since transition as the temporal dimension.¹⁰ To measure the severity of lustration in a given country k , we constructed a variable $severity_k$, calculated as the total number of positive transitional events divided the total number of transitional events plus one.¹¹ This variable takes on the value of 0 when country k did not attempt any transitional event at all. The number of negative transitional events is strictly lower than the number of positive transitional events, ensuring that a state with any transitional justice event at all will have a non-zero score. This ratio approaches 1 when the ratio of positive events to total events increases. Note that, whereas the dependent variable, programmaticness, is measured at the party-level, severity is measured at the country-level. The addition of 1 in the denominator of the measure ensures that the denominator is different than zero.

Hypothesis 1: As the severity of lustration legislation increases, all else kept constant, blackmail will be less successful and the quality of representation will increase.

A second empirical implication of the model is that as the costs of revealing the dissident’s skeletons increase, the degree of programmatic representation should decrease. In the context of democracies recovering from authoritarian rule, revealing that a dissident collaborated in the past with the authoritarian regime can be damaging for his political

¹⁰Each country was coded independently by two coders and the inter-coder reliability reached consistently over 90%. To arrive at the totals we pooled transitional justice events from both coding sources.

¹¹ $severity_k = \frac{Positive_events_k}{Total_events_k+1}$

reputation. Therefore, the costs of revealing skeletons, amount to the electoral costs of being fired (by the voters). However, voters may care to various degrees about revelations of evidence produced by the agent and their preference for a specific candidate may more or less directly translate into a candidate's chances of remaining in office. For instance, voters from younger democracies may care more about the authoritarian past compared to voters in older ones (Stan 2006, Bernhard & Kubik 2014, Wilde 1999, Cohen 1995, De Brito, Enríquez & Aguilar 2001). This claim is supported by literature on historical memory (Bernhard & Kubik 2014), but is also true by the mere fact that a larger share of voters in a young democracy will have experienced (and can remember) life under autocracy (Pop-Eleches 2007, Pop-Eleches & Tucker 2011, Pop-Eleches & Tucker 2012). In electoral systems with the personal vote (such as SMD or OLPR), these costs could be higher than in systems that do not give voters a say over which specific candidate gets elected into office (Carey & Shugart 1995). Thus the "personal vote" could serve as an alternative measure of the costs to the dissident of having embarrassing evidence revealed. A third way to operationalize this parameter would be to look at turnout in the first free elections following the transition to democracy. We offer a coarse operationalization for cost of firing: the time lapsed since the transition. We argue that, the longer time has passed, voters might be less likely to care about revelations of skeletons in the closet.

Hypothesis 2: As the time lapsed since the transition to democracy increases, all other things held constant, blackmail will be less effective and the degree of programmatic representation will increase.

Thirdly, a somewhat unexpected implication of our formal model is that the ideological distance between the dissident and his agent has no effect on the degree of programmatic representation. We operationalize the distance in ideal points between the dissident and the agent as the ideological distance (left-right scale) as measured by the DALP expert surveys

between the successor authoritarian party and each other party.

Hypothesis 3: Programmatic representation should be constant in the ideological distance between a country’s successor authoritarian party and a non-successor party.

3.1 Statistical Models

The way we propose to operationalize the implications of the model implies a hierarchical structure to the data. While both the dependent variable, programmaticness, and ideological distance are party-level attributes, the remaining variables are measured at the country-level. Because parties within countries are not independent from each other, an ordinary least squares regression (OLS) risks biased and inefficient estimation. Traditional solutions for this (fixed effects or clustered standard errors) are inapplicable because some countries only have one party that is not successor authoritarian. Thus, instead of OLS, we use a multilevel approach. This approach structures the variation in the dependent variable according to different levels, and can be thought of as a compromise between assuming all parties are independent from each other with no relevant country-level characteristics, and assuming all variation across parties can be explained by the country that they belong to.¹² The full model is expressed here:

$$\begin{aligned} \text{cosal_3}_{p,k} &= \beta_k + \beta_d * d_party_{p,k} + \beta_p * severity_k + \beta_F * n_year + \sum X_k + e_{p,k} \quad (1) \\ \beta_k &\sim \mathcal{N}(\mu_{\beta_k}, \sigma_{\beta_k}^2) \end{aligned}$$

As can be seen, the model estimates the average effect of all the variables of interest

¹²The country-level coefficients are assumed to share an underlying distribution with a global mean (μ_{β_k}) and standard deviation (σ_{β_k}). This is contrary to the traditional fixed-effects approach, which would include one intercept per country and assume these intercepts to be independent of each other. The advantage of our multilevel model is that when there is very little observed variation at the country-level, the coefficient can be estimated to be close to the mean μ_{β_k} .

in a manner similar to regular OLS. This means that we estimate the average effect of the variables of interest on the degree of programaticness of the dissident party. However, we acknowledge that there might be country-specific characteristics that explain the remaining variation, and we estimate that using a series of country-level intercepts (β_k). The results of the estimated models are reported in the next table:

Table 1: Explaining programaticness (country intercepts not shown)

	Model 1	Model 2	Model 3	Model 4	Model 5
<i>d_party</i>	-0.008 (0.005)	-0.008 (0.005)	-0.008 (0.005)	-0.008 (0.005)	-0.009 (0.005)
<i>severity</i>		0.254*** (0.090)	0.257*** (0.090)	0.212** (0.089)	0.226** (0.091)
<i>n_year</i>			-0.004 (0.002)	-0.003 (0.002)	-0.002 (0.002)
<i>press_freedom</i>				0.005** (0.003)	0.005* (0.003)
<i>opp_status</i>					0.046 (0.053)
<i>no_succ</i>					-0.038 (0.064)
Constant	0.367*** (0.031)	0.238*** (0.054)	0.335*** (0.084)	-0.009 (0.190)	-0.018 (0.209)
Observations	207	207	207	207	207
Log Likelihood	110.024	112.209	108.196	105.097	101.761
Akaike Inf. Crit.	-212.048	-214.418	-204.392	-196.195	-185.522
Bayesian Inf. Crit.	-198.717	-197.754	-184.396	-172.866	-155.528

Note:

*p<0.1; **p<0.05; ***p<0.01

The resulting estimations are consistent with the implications of our theoretical model. Specifically, we find that the ideological distance between the dissident and the successor

party is not statistically different than zero. This supports the somewhat surprising finding from the theoretical model that the level of programmaticness does not depend on the distance between the ideal points of the successor authoritarian party and the non-successor parties. This lack of statistical significance of the coefficient associated with *d_party* is robust across all models.

We address two possible concerns regarding this lack of statistical significance. The first is that this zero coefficient is driven by preexisting closeness between the opposition and the authoritarian parties. That is, it could be that parties that existed during the authoritarian period are, at the same time, ideologically closer to the authoritarian parties and more likely to be blackmailed. We control for this possible confounder using a measure of how active was the opposition during the authoritarian era. The data provided by Cheibub, Gandhi & Vreeland (2010) report whether the authoritarian regime had no legislature, or a single or multiparty legislature. In our sample, only two authoritarian regimes had no legislature at all, so we turned the variable *opp_status* into an indicator taking the value of 1 if there were multiple parties represented in the legislature, and 0 otherwise. As can be seen in Model 5, the inclusion of this variable does not change our results.

The second concern is related to our handling of missing values. That is, not all countries have a successor authoritarian party. In these instances, we inputted the average distance between successor authoritarian parties and the rest of the parties in those countries that do have a successor party. Including an indicator variable (*no_succ*) that takes the value of 1 if there is no missing successor allows us to verify if the estimate of the coefficient for *d_party* is biased. The coefficient for this indicator is not statistically significant nor does it change the point estimates for the other variables.¹³ This further supports our hypothesis that variation in ideological proximity of authoritarian parties and parties under democracy does not explain how programmatic are parties in general.

The estimations also support the prediction that *severity* of lustration measures would

¹³See the Appendix for alternative treatment of missing values of successor parties.

impact positively the programmaticness of parties. The effect is statistically significant at the 0.05% confidence level in all specifications. In addition, its average effect is quite sizeable: going from no attempts of transitional justice at all (a *severity* score of 0) to a score of 0.9 (the largest score on the data) increases programmaticness an average of 0.2 points. Since programmaticness ranges from 0 to 1, this effect is not minor.

Models 3 and 4 include the variable *n_year*, which captures time since transition to democracy as a proxy for the cost of firing (F). As can be seen, the coefficient is negatively associated with the programmaticness of a party, which supports the idea that the longer time has passed since transition, voters might care less about their alleged collaborations and hence blackmail might be less effective. However, the coefficient is not statistically significant in either model, something that we attribute to the inaccuracy of the proxy. Future iterations will try to incorporate different operationalizations.

Finally, models 4 and 5 include the variable *press_freedom*, which is our way of operationalizing parameter c , cost of bluffing. Recall our model predicts that an increase in the costs of bluffing to the agent does not increase the degree of programmatic representation, but that all things held constant, it increases the chances of an equilibrium where transitional justice does not work as it ought to. Hence, there is no direct effect of the cost of bluffing on programmatic representation, though there is an effect on whether politicians are bluffed or blackmailed. We include press freedom, measured with the rating provided by Reporters Without Borders (Press Freedom)¹⁴ as a way to evaluate the performance of the remaining variables controlling for “costs of bluffing.” If our theoretical model is correct, we would not expect the inclusion of this variable to change our findings.

The intuition behind operationalizing c with press freedom is as follows: When an agent tries to bluff the dissident into thinking that the agent is in possession of evidence against the politician, such a bluff can be easily exposed in an environment with free media. Indeed, a wide literature discusses how journalists are on par with state-sponsored archives (Laplante

¹⁴To avoid confusion, the variable was linearly transformed so that lower rankings mean lower freedom of the press in a given country.

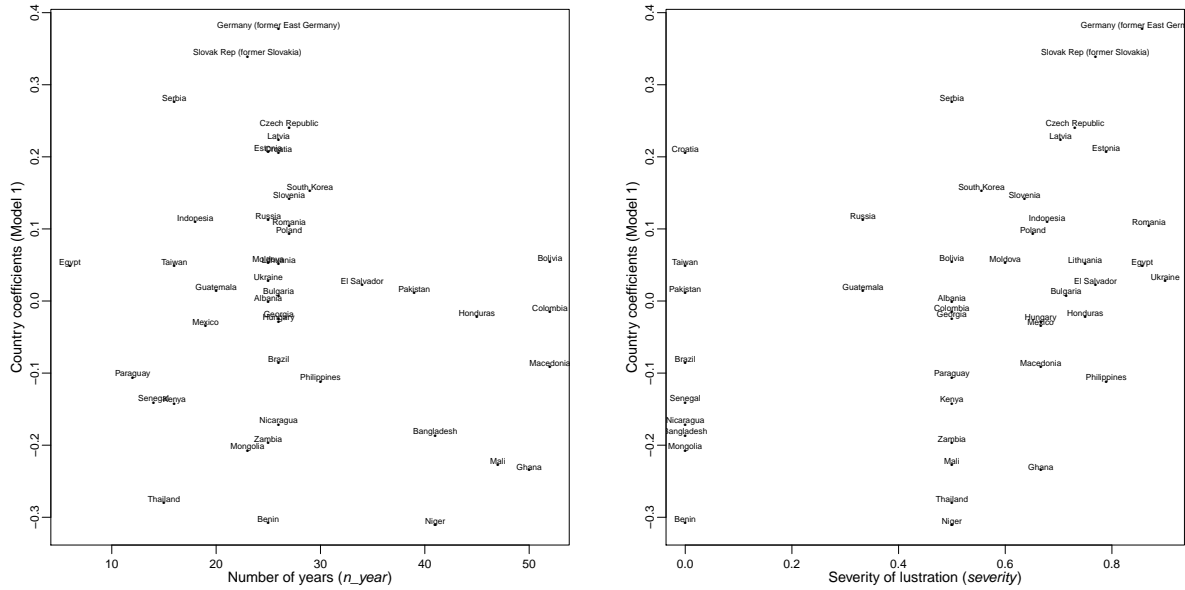
& Phenicie 2009, Pinto 2010, Chapman 2009) when it comes to revealing evidence of former secret police conduct. Although such revelation is not subject to the same level of scrutiny as state sponsored legislation, a high degree of media independence would lead to the bluff being called, whereas the media's failure to corroborate the claim that the dissident in question was a former collaborator would weaken the agent's bluff. Models 4 and 5 incorporate this variable and show that, although this variable is positively related to the degree of programaticness of a political party, it does not change the estimated effects of the rest of the variables. Note, however, that the constant term is no longer statistically significant when *pressfreedom* is included. We interpret this as a type of omitted variable bias: in models 1, 2, and 3 the average effect of a free press is included in this constant.

Finally, the next figures provide a visual representation of the relationship between the relevant independent variables at the country level and the average programaticness of parties per country, as estimated in model 1. Model 1 included only the regressor that was measured at the party level: ideological distance, and estimated a series of country-specific intercepts which capture the average country-level programaticness. As can be seen, the left panel shows no effect between number of years (horizontal axis) and average country programaticness. This is consistent with the results presented, but the figure suggests that this lack of statistically significant relationship is not due to misspecification of functional form (that is, we do not observe any specific pattern in the figure).

The panel on the right illustrate the positive relationship between severity of lustration (horizontal axis) and average programaticness (vertical axis). Furthermore, the panel shows that this pattern is much stronger for the observations that fall on the right half of the panel. In sum, this figure illustrates the relationship posited here.

As the plot shows, the relationship between the country-level variation in programaticness of the political parties is stronger with *severity* than with *n_year*. These findings support our main empirical and theoretical arguments. Empirically, there is reason to believe that transitional justice is better measured as a contested process, rather than simply indicating

Figure 4: Country intercepts (Model 1) and country-level variables



the adoption, or not, of these measures.

Theoretically, the empirical evidence presented here shows, firstly, that the ideological distance between authoritarian and successor parties do not shape the overall degree of programmaticness. This is surprising given that one would expect that parties that are closer to the successor might be more willing to give into blackmail than parties that are farther apart. This is so because the policies politicians are being asked to implement are closer do not depart that much from their ideal points. Yet we find no evidence of this, as the coefficient of d_{party} is not statistically significant.

Secondly, the evidence shows that the severity of transitional justice is positively related to the degree of programmaticness. This result is promising insofar as it suggests that lustration policies produce, on average, more programmatic parties after transition. However, the theoretical model also shows that lustration by itself does not prevent blackmail of former dissidents turned politicians. In order for the proper functioning of lustration policies, this must come accompanied by a free press that will prevent bluffing.

4 Conclusion

In this paper we investigated how policies set up in new democracies for dealing with authoritarian elites shape the quality of linkages forged between voters and the parties they vote for in post-authoritarian states.

We used formal theory to model how members of the former enforcement apparatus can use blackmail with secret police files to pressure new democratic elites into implementing policies that are amenable to former authoritarian elites instead of policies representing voters' preferences. Theoretically, we find that former authoritarian elites' ability to shape policy is higher when transitional justice is legislated less frequently and that their influence decreases with how much the voters care for skeletons of the authoritarian past in their politicians' closet. Also, to our surprise and in contrast to what most of the literature has been positing, the ideological distance between the former autocrats and the former dissidents has no bearing on the ability of former autocrats to extract concessions. Thus, the polarized policy recommendations found in the transitional justice literature recommending dealing with former autocrats either by banishing them from politics entirely or by integrating them into the political class are both theoretically unfounded.

We evaluate the empirical implications of the model with an original dataset that includes a time series of all personnel transitional justice events in former authoritarian party-based regimes.

Our dataset was sufficient to confirm especially this first theoretical point. We also find that holding all else constant, lack of free press can sabotage transitional justice efforts as the absence of independent media allows blackmail based on bluffing, even when no actual evidence exists.

To date, there has been very little research on the effect of transitional justice processes on the long-term quality of representation. Nobody has examined how transitional justice affects the survival of former authoritarian networks and the ability of former authoritarian elites to leverage the secret information assembled under their tenure to have enduring impact

on the policies of new democracies.

By linking transitional justice to the nature of secret authoritarian legacies we contribute to a fuller understanding of the relationship between transitional justice, rule of law and the quality of representation in new democracies. For instance, our theory allows us to reinterpret some dominant political narratives about the effects of delayed transitional justice on the quality of representation. In the past, scholars have criticized countries such as Poland and Hungary for being late comers to transitional justice. The initially implemented transitional justice policies in these two countries were extremely mild leading many transitional justice experts to believe that “living well was the best revenge.” However, although they started off the democratization period with very little lustration activity, gradually they have acquired quite a substantial body of transitional justice legislation and today their party systems, particularly Poland’s stand out as highly representative of the general public. In light of our argument presented here, transitional justice policies implemented later after the transition, by decreasing the costs of firing politicians, make possible equilibria that lead to higher quality representation than hastily implemented lustration laws. This is because in the immediate aftermath of transition to democracy, when the risks of having politicians’ skeletons in the closet exposed are highest, former secret police agents have the highest leverage to make their lustration blackmail effective even when they are only bluffing.

Paradoxically, the less the public cares for skeletons, the better the transitional justice policy works, because it makes bluffing less frequent.

From a normative point of view, we might feel the public should shame politicians who had a murky past, but our model suggests that you cannot have both shaming in effect and a transitional justice policy that works as it should.

References

- Bernhard, M. & Kubik, J. (2014), *Twenty Years After Communism: The Politics of Memory and Commemoration*, Oxford University Press.
- Brun, D. A. & Diamond, L. (2014), *Clientelism, social policy, and the quality of democracy*, JHU Press.
- Carey, J. M. & Shugart, M. S. (1995), 'Incentives to cultivate a personal vote: A rank ordering of electoral formulas', *Electoral studies* **14**(4), 417–439.
- Casper, B. A. & Tyson, S. A. (2014), 'Military Compliance and the Efficacy of Mass Killings at Deterring Rebellion', *Available at SSRN 2434297*.
- Chapman, A. R. (2009), 'Truth finding in the transitional justice process', *Assessing the impact of transitional justice: challenges for empirical research* pp. 91–114.
- Cheibub, J. A., Gandhi, J. & Vreeland, J. R. (2010), 'Democracy and dictatorship revisited', *Public Choice* **143**(1-2), 67–101.
- Closa Montero, C. (2010), 'Study on how the memory of crimes committed by totalitarian regimes in europe is dealt with in the member states'.
- Cohen, S. (1995), 'State crimes of previous regimes: Knowledge, accountability, and the policing of the past', *Law & Social Inquiry* **20**(1), 7–50.
- David, R. (2003), 'Lustration laws in action: the motives and evaluation of lustration policy in the czech republic and poland (1989–2001)', *Law & Social Inquiry* **28**(2), 387–439.
- David, R. (2006), 'From Prague to Baghdad: Lustration Systems and their Political Effects1', *Government and Opposition* **41**(3), 347–372.
- David, R. (2011), *Lustration and transitional justice: Personnel systems in the Czech Republic, Hungary, and Poland*, University of Pennsylvania Press.
- De Brito, A. B., Enríquez, C. G. & Aguilar, P. (2001), *The Politics of Memory and Democratization: Transitional Justice in Democratizing Societies: Transitional Justice in Democratizing Societies*, Oxford University Press.
- Ellis, M. S. (1996), 'Purging the past: The current state of lustration laws in the former communist bloc', *Law and Contemporary Problems* **59**(4), 181–196.
- Elster, J. (1998), 'Coming to terms with the past. a framework for the study of justice in the transition to democracy', *European Journal of Sociology* **39**(01), 7–48.
- Elster, J. (2004), *Closing the books: Transitional justice in historical perspective*, Cambridge University Press.
- Grzymala-Busse, A. M. (2002), *Redeeming the communist past: The regeneration of communist parties in East Central Europe*, Cambridge University Press.

- Halmai, G., Scheppele, K. L. & McAdams, A. J. (1997), *Transitional Justice and the Rule of Law in New Democracies*, University of Notre Dame Press Notre Dame, IN.
- Holmes, S. (1994), 'End of decommunization, the', *E. Eur. Const. Rev.* **3**, 33.
- Horne, C. M. (2009), 'International legal rulings on lustration policies in central and eastern europe: Rule of law in historical context', *Law & Social Inquiry* **34**(3), 713–744.
- Horne, C. M. (n.d.), *Building Trust and Democracy in Transition: Assessing the Impact of Transitional Justice in Post-Communist Countries*, Oxford University Press.
- Kaminski, M. M., Nalepa, M. & O'Neill, B. (2006), 'Normative and strategic aspects of transitional justice', *Journal of Conflict Resolution* pp. 295–302.
- Kitschelt, H. (1999), *Post-communist party systems: competition, representation, and inter-party cooperation*, Cambridge University Press.
- Kitschelt, H. & Freeze, K. (2010), Programmatic party system structuration: Developing and comparing cross-national and cross-party measures with a new global data set, in 'APSA Annual Meeting, Washington, DC'.
- Kitschelt, H., Freeze, K., Kolev, K. & Wang, Y.-T. (2009), 'Measuring democratic accountability: An initial report on an emerging data set', *Revista de Ciencia Política* **29**(3), 741–773.
- Kritz, N. J. (1995), *Transitional justice: how emerging democracies reckon with former regimes*, Vol. 2, US Institute of Peace Press.
- Laplante, L. J. & Phenicie, K. (2009), 'Mediating post-conflict dialogue: the media's role in transitional justice processes'.
- Letki, N. (2002), 'Lustration and democratisation in east-central europe', *Europe-Asia Studies* **54**(4), 529–552.
- Lundy, P. & McGovern, M. (2008), 'Whose justice? Rethinking transitional justice from the bottom up', *Journal of Law and Society* **35**(2), 265–292.
- Mallinder, L. (2008), *Amnesty, Human Rights and Political Transitions: Bridging the Peace and Justice Divide*, Hart Publishing Limited.
- Mayer-Rieckh, A. & De Greiff, P. (2007), *Justice as prevention: vetting public employees in transitional societies*, Social Science Research Council.
- Nalepa, M. (2008), 'To punish the guilty and protect the innocent comparing truth revelation procedures', *Journal of Theoretical Politics* **20**(2), 221–245.
- Nalepa, M. (2010), *Skeletons in the closet: Transitional justice in post-communist Europe*, Cambridge University Press.

- Nedelsky, N. (2013), 'From velvet revolution to velvet justice: The case of slovakia', *After Oppression* pp. 390–417.
- Olsen, T. D. (2010), 'Transitional Justice In Balance: Comparing Processes, Weighing Efficacy Author: Tricia D. Olsen, Leigh A. Payne, Andrew G.'.
- Pettai, E.-C. & Pettai, V. (2014), *Transitional and retrospective justice in the Baltic States*, Cambridge University Press.
- Pinto, A. C. (2010), 'The authoritarian past and south european democracies: An introduction', *South European Society and Politics* **15**(3), 339–358.
- Pop-Eleches, G. (2007), 'Historical legacies and post-communist regime change', *Journal of Politics* **69**(4), 908–926.
- Pop-Eleches, G. & Tucker, J. A. (2011), 'Communism's shadow: postcommunist legacies, values, and behavior', *Comparative Politics* **43**(4), 379–408.
- Pop-Eleches, G. & Tucker, J. A. (2012), 'Associated with the past? communist legacies and civic participation in post-communist countries', *East European Politics & Societies* p. 0888325412465087.
- Ritter, E. H. & Wolford, S. (2012), 'Bargaining and the effectiveness of international criminal regimes', *Journal of Theoretical Politics* **24**(2), 149–171.
- Slater, D. & Wong, J. (2013), 'The strength to concede: ruling parties and democratization in developmental asia', *Perspectives on Politics* **11**(03), 717–733.
- Stan, L. (2006), 'The vanishing truth? politics and memory in post-communist europe', *East European Quarterly* **40**(4), 383–409.
- Stan, L. (2012), 'Witch-hunt or moral rebirth? romanian parliamentary debates on lustration', *East European Politics & Societies* **26**(2), 274–295.
- Stan, L. (2013), 'Reckoning with the communist past in romania: A scorecard', *Europe-Asia Studies* **65**(1), 127–146.
- Stan, L. & Nedelsky, N. (2015), *Post-communist Transitional Justice: Lessons from Twenty-five Years of Experience*, Cambridge University Press.
- Stan, L. & others (2009), *Transitional justice in Eastern Europe and the former Soviet Union: Reckoning with the communist past*, Routledge.
- Thoms, O. N., Ron, J. & Paris, R. (2008), 'The effects of transitional justice mechanisms: A summary of empirical research findings and implications for analysts and practitioners', *Ottawa: Centre for International Policy Studies* .
- Van der Merwe, H., Baxter, V. & Chapman, A. R. (2009), *Assessing the impact of transitional justice: Challenges for empirical research*, US Institute of Peace Press.

Volčič, Z. & Simić, O. (2013), Localizing transitional justice: Civil society practices and initiatives in the balkans, *in* 'Transitional Justice and Civil Society in the Balkans', Springer, pp. 1–14.

Wilde, A. (1999), 'Irruptions of memory: expressive politics in chile's transition to democracy', *Journal of Latin American Studies* **31**(02), 473–500.

5 Formal Appendix

5.1 Equilibrium Analysis

We first will verify the plausibility of the pure separating equilibrium in which all agents equipped with evidence choose a different strategy from all agents not equipped with such evidence.

5.1.1 Pure separating equilibrium

Since it is never rational for A to not make a demand when he has evidence against the dissident,¹⁵ the only possible separating equilibrium is: $(0, R, 1, x^*)$. The requirement for this to be a Bayesian equilibrium is that the posterior beliefs are $Pr(E|0) = 0, Pr(\sim E|0) = 1, Pr(E|1) = 1, Pr(\sim E|1) = 0$. The following four steps will lead to uncovering the conditions for this equilibrium.

1. First, suppose x^* is the proposal accepted in equilibrium. D knows that if he proposes $x < x^*$, he will be fired with certainty. Therefore, on the one hand, if he is going to make a proposal, in equilibrium, it must be accepted, because if it were to be rejected D should propose $x = d$. On the other hand, if a demand is placed and D makes a counteroffer $x \neq d$, D 's utility must be greater from having his proposal accepted than from having it rejected, i.e.,

$EU_D(x|0, 1) \geq EU_D(d|0, 1)$, which is equivalent to $-|d - x| \geq -|d - d| - F$. The last expression simplifies to

$$x \geq d - F$$

2. Second, note, that the lowest x^* that the dissident should be willing to accept is $x^* = d - F$. Therefore, A 's optimal rejection region is $R = (d - F, d]$
3. Third, note that to prevent A from bluffing (and making demands when there is no evidence), it has to be the case that that $EU_A(0|x^*, \sim E) \leq EU_A(1|x^*, \sim E)$, which is equivalent to $-|a - x^*| - c \leq -|d - a|$. This last expression simplifies to $x^* \geq d - c$, which after substituting x^* from above gives:

$$F \leq c$$

4. Fourth, for x^* to be a feasible proposal it has to be the case that $x^* \geq 0$, that is $d - F \geq 0$ otherwise D should just propose 0. Given the assumption that $F < d$, this always holds.

We conclude that a pure separating equilibrium exists only when the cost of being fired (F) relative to the cost of bluffing c is quite low. In this pure separating equilibrium,

¹⁵Note that D cannot make a counterproposal x when A has not made a demand. Thus, the worst A can do when making a demand is $-|a - d|$, which would be his payoff if D 's counterproposal were $x = a$, that is if D made no concession at all. But without making any demand A is guaranteed to receive $-|a - d|$ and no more.

blackmail is effective with probability p and its magnitude $(d - F)$ is directly proportional to the cost of firing. In the next two sections, we find how this departure from programmatic representation compares with the effectiveness of blackmail under the pure pooling and hybrid equilibria.

5.1.2 Pure pooling equilibrium

In pooling equilibria, agents with and without evidence will choose the same action, implying that the dissident cannot update his prior beliefs to posterior beliefs, conditioning on the agent's action. In the analysis above, we established that the agent will never refuse to place a demand when evidence is present. Thus, the only possibility of a pooling equilibrium in this game is $(0, R, 0; x'')$ with accompanying beliefs: $Pr(E|0) = p, Pr(\sim E|0) = (1 - p)$ ¹⁶. In this equilibrium, the agent always places a demand and the the dissident always offers the same counterproposal, x'' .

1. For such an equilibrium to hold, the dissident has to prefer to have his proposal accepted to being fired (in which case, he would simply propose his ideal point, d). Thus, it must be the case that $EU_D(x|0, 0) \geq EU_D(d|0, 0)$, which is equivalent to $-|d - x| \geq p(-|d - d| - F) + (1 - p) * 0$.¹⁷ The last expression simplifies to :

$$x \geq d - pF$$

2. Since $x'' = d - pF$ is the lowest proposal the dissident will accept, given his beliefs, the agent's optimal rejection region is $(d - pF, d]$.
3. To ensure the agent always has an incentive to place a demand, it has to be the case that $EU_A(0|x'', \sim E) \geq EU_A(1|x'', \sim E)$, which is equivalent to $-|a - x''| - c \geq -|d - a|$. This last expression simplifies to $x'' \geq d - c$, which after substituting x'' from above gives

$$F \geq \frac{c}{p}.$$

4. Finally, as before, to be feasible, x'' has to lie between a and d , i.e., $a < d - pF < d$., the right hand side of which is again ensured by a positive F and p . The left hand side of the inequality is ensured by our assumption $d < F$.

Summing up, a pure pooling equilibrium exists only when the cost of being fired (F) relative to the cost of bluffing c is quite high. In this pure pooling equilibrium, lustration blackmail is always effective (takes place with probability 1). the distortion it causes relative to the dissident's ideal point and is pF . It is directly proportional to the cost of firing and the extent to which evidence exists. Our final subsection of the equilibrium analysis looks at the effectiveness of lustration blackmail under the semi-pooling (hybrid) equilibrium.

¹⁶Note that paths that involve the agent not making a demand are off the equilibrium path and we do not have to specify the beliefs there

¹⁷Note that since the dissident cannot tell which type—with or without evidence—he is facing any better than he could before the agent took an action, his expected utility from making a proposal outside of the acceptance region is weighted by his priors about the probability that evidence exists.

5.1.3 Semi-separating equilibrium

In addition to the pure separating equilibrium discussed above, we also derive the conditions (and verify their plausibility) of a semi-pooling or (or hybrid) equilibrium. In this equilibrium, the agent plays a mixed strategy. He always makes a demand when evidence is present, but he also with some probability λ makes a demand if evidence does not exist (and with probability $1 - \lambda$ does not make a demand). Consequently, any hybrid equilibrium must fit the format $(1, R, \lambda; x')$.¹⁸ Note the beliefs consistent with this semi-pooling equilibrium are: $Pr(E|0) = 0, Pr(\sim E|1) = \lambda, Pr(\sim E|0) = 1, Pr(E|0) = 1 - \lambda$. These beliefs will be used in the calculation of the expected utilities. We proceed in six steps.

1. First, to find the equilibrium value of λ^* , we calculate the expected utility of the dissident from responding x to the agent's demand ($EU_D(x'|0, \lambda)$) and set it equal to the expected utility of the dissident's choosing an x which is outside of the agent's acceptance region ($EU_D(d|0, \lambda)$). This yields the equality $p(-|d - x'|) + (1 - p)\lambda(-|d - x'|) + 0 = p(-|d - d| - F) + (1 - p)\lambda(-|d - d| + 0 = -pF)$
2. Next, assuming that x' is the equilibrium proposal that falls into A 's acceptance region we require that $p(-|d - x'|) + (1 - p)\lambda(-|d - x'|) = -pF$, which in terms of λ can be stated as:

$$\lambda = \frac{p(d - x') - pF}{(1 - p)[- (d - x')]} \quad (2)$$

3. To ensure that $0 < \lambda < 1$ and is a probability F must satisfy:

$$d - x' < F < \frac{d - x'}{p} \quad (3)$$

4. Next, to pin down x' , we make use of the fact that when evidence does not exist, A must be indifferent between placing a demand and not placing one, i.e.: $EU_A(1|x', \sim E) = EU_A(0|x', \sim E)$, which reduces to: $-|d - a| = |-a - x'| - c$. the last equality, can be written in terms of x' as $x' = d - c$.
5. Finally, substituting x' into equation 2:

$$\lambda = \frac{p(F - c)}{(1 - p)c} \quad (4)$$

6. To get scope conditions for the semi-pooling equilibrium, we can substitute x' into condition 3:

$$c < F < cp$$

And to ensure that it is between a and d , we need $d - c > a = 0$.

¹⁸This means that the agent makes a demand with probability 1 if evidence exists and with probability λ if evidence does not exist.

6 Appendix: Empirics

Our dependent variable consists of a score at the party-level. Table 2 shows countries included in the sample, and the number of parties per country.

Table 2: Countries and number of parties per country included

Albania (4), Bangladesh (3), Benin (5), Bolivia (4) , Brazil (9), Bulgaria (6), Colombia (7), Croatia (9), Czech Republic (4), Egypt (4), El Salvador (2) , Estonia (6) , Georgia (4), Germany (former East Germany) (6), Ghana (2), Guatemala (6), Honduras (1), Hungary (4), Indonesia (6), Kenya (3), Latvia (7), Lithuania (5), Macedonia (4), Mali (5), Mexico (4) , Moldova (5), Mongolia (4), Nicaragua (2), Niger (3), Pakistan (6), Paraguay (6), Philippines (7), Poland (5), Romania (5), Russia (5), Senegal (4), Serbia (6), Slovak Rep (former Slovakia) (5), Slovenia (7), South Korea (2), Taiwan (1), Thailand (7), Ukraine (4) , Zambia (4).
--

We build from the original dependent variable *cosalpo_4* reported by DALP. We followed the author’s instructions and created our own cohesion and salience measure. Plot ?? shows two alternative measures that we propose, and compares them to the original variable provided by the authors. The first variable that we propose, *cosal_3* does not include the placement of the parties on a left to right scale as a relevant item (question *dw*). The second one (*cosal_4*) does. The analysis conducted here used *cosal_3* as a dependent variable, and this choice was theoretical. Given that we rely on the *dw* question to obtain the variable *d_party*, we considered the prudent approach of not including it in the final analysis.

7 Appendix: Robustness Checks

We present three robustness checks for our treatment of missing successor parties. The first robustness check, reported in table 3, substitutes missing ideological placement of successor parties with the average of the non-successor parties within the country. The second approach substitutes missing values with the average ideological placement of successor parties across all countries. As can be seen from the models on table 4 , this approach does not change most of the substantive findings, but the variable *d_party* is statistically significant and positive. Finally, we conduct the same estimations but only on the subset of countries with successor parties. Table 5 includes these cases. Our substantive findings do not change, although in these last set of models we do find support for the effect of *n_year* in average party programaticness.

Table 3: Explaining programaticness (country intercepts not shown)

	Missing successor as average of parties in country				
	Model 1a	Model 2a	Model 3a	Model 4a	Model 5a
<i>d_party</i>	-0.008 (0.005)	-0.008 (0.005)	-0.008 (0.005)	-0.008 (0.005)	-0.009 (0.005)
<i>severity</i>		0.254*** (0.090)	0.257*** (0.090)	0.212** (0.089)	0.226** (0.091)
<i>n_year</i>			-0.004 (0.002)	-0.003 (0.002)	-0.002 (0.002)
<i>press_freedom</i>				0.005** (0.003)	0.005* (0.003)
<i>opp_status</i>					0.046 (0.053)
<i>no_succ</i>					-0.038 (0.064)
Constant	0.367*** (0.031)	0.238*** (0.054)	0.335*** (0.084)	-0.009 (0.190)	-0.018 (0.209)
Observations	207	207	207	207	207
Log Likelihood	110.024	112.209	108.196	105.097	101.761
Akaike Inf. Crit.	-212.048	-214.418	-204.392	-196.195	-185.522
Bayesian Inf. Crit.	-198.717	-197.754	-184.396	-172.866	-155.528

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 4: Explaining programaticness (country intercepts not shown)

	Missing successor as average of successors across countries				
	Model 1b	Model 2b	Model 3b	Model 4b	Model 5b
<i>d_party</i>	-0.008 (0.005)	-0.008 (0.005)	-0.008 (0.005)	-0.008 (0.005)	-0.009 (0.005)
<i>severity</i>		0.254*** (0.090)	0.257*** (0.090)	0.212** (0.089)	0.226** (0.091)
<i>n_year</i>			-0.004 (0.002)	-0.003 (0.002)	-0.002 (0.002)
<i>press_freedom</i>				0.005** (0.003)	0.005* (0.003)
<i>opp_status</i>					0.046 (0.053)
<i>no_succ</i>					-0.038 (0.064)
Constant	0.367*** (0.031)	0.238*** (0.054)	0.335*** (0.084)	-0.009 (0.190)	-0.018 (0.209)
Observations	207	207	207	207	207
Log Likelihood	110.024	112.209	108.196	105.097	101.761
Akaike Inf. Crit.	-212.048	-214.418	-204.392	-196.195	-185.522
Bayesian Inf. Crit.	-198.717	-197.754	-184.396	-172.866	-155.528

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 5: Explaining programaticness (country intercepts not shown)

	Only countries with successor parties included			
	Model 1c	Model 2c	Model 3c	Model 4c
<i>d_party</i>	-0.008 (0.005)	-0.008 (0.005)	-0.008 (0.005)	-0.009* (0.005)
<i>severity</i>		0.250** (0.103)	0.292*** (0.098)	0.239** (0.099)
<i>n_year</i>			-0.007** (0.003)	-0.007** (0.003)
<i>press_freedom</i>				0.006* (0.003)
<i>opp_status</i>	0.388*** (0.035)	0.261*** (0.061)	0.437*** (0.094)	0.039 (0.239)
Observations	152	152	152	152
Log Likelihood	81.820	83.230	81.099	77.850
Akaike Inf. Crit.	-155.639	-156.460	-150.197	-141.701
Bayesian Inf. Crit.	-143.544	-141.341	-132.054	-120.534

Note:

*p<0.1; **p<0.05; ***p<0.01