Can Transitional Justice Improve the Quality of Representation in New Democracies?

Milena Ang
University of Chicago

Monika Nalepa
University of Chicago

Word count: 10363

Abstract

Can transitional justice enhance democratic representation in countries recovering from authoritarian rule? We argue that lustration, the policy of revealing secret collaboration with the authoritarian regime, can prevent former authoritarian elites from extorting policy concessions from past collaborators who have become elected politicians. Absent lustration, former elites can threaten to reveal information about past collaboration unless politicians implement policies these elites desire. In this way, lustration laws enable politicians to avoid blackmail and become responsive to their constituents, improving the quality of representation. We show that whether lustration enhances representation depends on its severity and the extent to which dissidents-turned-politicians suffer if their skeletons come out. We also find that the potential to blackmail politicians increases as the ideological distance between authoritarian elites and politicians decreases. We test this theory with an original personnel transitional justice dataset covering 84 countries that transitioned to democracy since 1946.

Acknowledgments: This research was supported by the National Science Foundation (Award Number 1658170) and the Center for Social Science Research Faculty Grant. The authors are grateful to Jaimie Bleck, Anna Grzymala-Busse, Karrie Koessel, Sue Stokes, Susanne Wengle, and three anonymous reviewers for their comments. We also thank Tiberiu Dragu, John Patty, and Scott Tyson for advice on the formal model. Participants of the Visions in Methodology Workshop at UC Davis in 2016 provided priceless comments and advice. The following research assistants contributed to compiling the original data: Genevieve Bates, Ipek Cinar, Ji Xue, Viivi Jarvi, Evgenia Olimpieva, Alexa Rosario, and Yonatan Litwin. All mistakes are the authors' responsibility.

1 Introduction

Does the presence of former authoritarian elites in the institutions of the new regime harm or hurt a new democracy's prospects for consolidation? On the one hand, former authoritarian elites can sabotage the transition using their expertise to capture resources previously controlled by the authoritarian state.[1] On the other, alienating these elites may force them to seek employment in criminal organizations presenting a challenge to the new democratic regime.[2] Given these opposing potential effects, it is not surprising that most empirical research has found ambiguous effects of personnel transitional justice, that is, of policies dealing with former autocrats and their collaborators in the aftermath of transition.

In this article, we address the relationship between personnel transitional justice and quality of democracy, focusing on the hidden dynamics between former authoritarian elites and politicians after transition. We begin by acknowledging that not all participants and collaborators of the authoritarian regime are equally threatening to democratic life, and that secret collaborators of the authoritarian regime are particularly destabilizing for democratic representation. This is so because authoritarian elites can blackmail them by threatening to reveal information about their involvement with the regime to extract policy concessions after transition. We also show that implementing personnel Transitional Justice, specifically lustration, can reduce this detrimental impact. We show that lustration weakens the ability of former authoritarian elites to leverage the secret information assembled under their tenure. Our explanation suggests that bringing the politicians' murky past to light can contribute to better quality of democratic representation.

Transitional Justice (TJ) comprises the "formal and informal procedures implemented

[1] Albertus and Menaldo 2014; Escribà-Folch and Wright 2015; Gandhi and Przeworski 2007.

[2] Lessing and Willis forthcoming.

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".[3] Transitional justice is built on four pillars: truth commissions, trials, victim compensation, and lustration. The latter, lustration, vets candidates for public office for ties to the former authoritarian secret police. Lustration opens up the archives of the former regime, and uncovers those who had collaborated with the secret police. Proven collaborators are then either explicitly banned from holding office or revealed as collaborators to the voters, who subsequently decide whether to cast their votes for these compromised politicians.

Lustration policies are frequently supported using *backward looking* arguments. For example, some claim that victims of espionage by the regime have the right to know who informed on them and their activities,[4] whereas others emphasize the importance of preventing former spies and their leading officers from playing key roles in public service.[5] However, in this paper we develop and present evidence for a *forward looking* argument for lustration: we suggest that revealing evidence of collaboration with the ancién regime prevents former authoritarian elites from influencing policy once transition has occurred. Where lustration is lacking, former authoritarian elites can pressure politicians into policy concessions by threatening to reveal information that could destroy politicians' careers. If blackmail is effective, we argue, politicians can lose their ability to represent voters. In this way, our theory formalizes a mechanism by which lustration policies can improve the quality of democratic representation after transition from authoritarian rule.

As a motivating illustration of this mechanism, consider the following anecdote from Poland. In February of 2016, the widow of Czesław Kiszczak, former Chief of the Communist Secret Police, discovered a thick secret police file containing evidence that Lech

[3]Kaminski, Nalepa, and O'Neill 2006.

[4]Stan 2012.

[5]Nedelsky 2013; Stan and others 2009; David 2011.

Wałęsa, former Nobel Peace Prize Laureate and anti-communist dissident leader, had served as a secret police collaborator between 1970 and 1976. In November 1990, Wałęsa was elected president for a five year term, which ended before Poland implemented a lustration program that would have vetted Wałęsa for connections to the communist secret police. 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 to question the quality of Wałęsa's presidency and the extent to which he represented interests of the electorate over those of the former secret police, and left others wondering about the prevalence of such acts of blackmail.

This anecdote suggests that former authoritarian elites can leverage secret information from politicians' past to blackmail these politicians. In turn, we argue that this blackmail hinders the quality of representation because politicians will implement policies desired by the former authoritarian elites, instead of the ones preferred by voters or the ones offered in campaign. Lustration, when implemented thoroughly, ought to enhance democratic representation by publicizing evidence of collaboration and making blackmail impossible. Our paper investigates this forward looking argument: that by uncovering the secrets that former autocrats could use in blackmail, lustration improves the quality of representation. [6]

[6]To be sure, not all political elites are as prominent as Lech Wałęsa, nor are they in charge of entire branches of government as Wałęsa was. Yet the impact of blackmailing individual former collaborators can be extremely destabilizing for a political party, particularly in parliamentary regimes, where the cohesiveness of parties depends on all members voting in unison (Carroll and Nalepa 2014; Carroll and Nalepa 2016). Because we are unable to collect minute data about the quality of representation of individual politicians, the empirical section focuses on the representativeness of parties (their ability to programmatically represent), thus in practice, our understanding of political elites is much more

We use a game theoretic model of incomplete information in order to reconstruct the blackmail mechanism. The model formalizes the way in which lustration limits the potential of blackmailing secret collaborators. A key strength of our model is that it isolates blackmail, or the potential of blackmail, as the only element that can impede the implementation of policies preferred by the politician. Since the politician is an elected representative, any deviations from her preferred policy is interpreted as hindering the quality of political representation. Hence, our game theoretic model captures the pathologies in representation that can only be due to blackmail, and it connects variation in the prevalence of blackmail to lack of lustration or poorly implemented lustration. In this way, we contribute to the growing literature of transitional justice and democratization by providing a theory of the effects of lustration on the quality of representation.

Our argument also contributes to the literature on a critical aspect of democratic stability: the linkages between citizens and political parties.[7] Scholars have noted that parties in newer democracies rely on strategies that range from programmatic party platforms to charismatic candidates or clientelism as a way to build linkages with citizens.[8] Mainwaring argues that under-institutionalized party systems, systems where parties are poorly organized and fragmented, and where electorates are volatile, can result in populist political elites who engage in clientelistic practices.[9] In young democracies, politicians may be more convincing at promising targeted benefits than policy, creating clientelistic ties with voters,[10] further hindering the establishment of programmatic parties. Political elites might also hold a monopoly on goods or services that the electorate wants to access, which also undermines representativeness by, for example, reducing the competitiveness of the elec-

rank and file than the Wałęsa anecdote suggests.

[7]Tavits 2005; Pop-Eleches 2010.

[8]Kitschelt and Wilkinson 2007.

[9]Mainwaring 1999.

[10]Keefer 2007.

tions.[11] Our argument contributes to this literature by identifying another mechanism that prevents the establishment of programmatic platforms in recently transitioned democracies: the blackmail of politicians by members of the previous authoritarian regime. To the extent that programmatic representation involves delivering policies that represent broad wishes of the electorate, our theory makes predictions about how transitional justice enables such representation by eliminating the possibility of blackmail.

Our final contribution is empirical. Despite a growing number of empirical studies examining the impact of lustration and transitional justice on democratic stability and peace,[12] 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 research on this topic concentrated on transitional justice as an outcome to be explained,[13] and only recently have scholars begun to investigate it as an independent variable. Among these recent contributions, research by Horne[14] and Olsen[15] is noteworthy because of their focus on lustration policies, although the data used

[11]Medina and Stokes 2007.

[12]Olsen 2010; Merwe, Baxter, and Chapman 2009; Thoms, Ron, and Paris 2008.

[13]Stan and others 2009; Lundy and McGovern 2008; Mallinder 2008; Pettai and Pettai 2014.

[14]In Horne 2017, the relationship between the nature of transitional justice mechanisms and trust is examined. The author suggests 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.

[15]Olsen 2010.

is somewhat limited: Olsen aggregates all lustration activity into a dichotomous variable indicating the presence or absence of lustration, and Horne includes only Post Communist countries.[16]

Our paper addresses these limitations by substantiating the theory using a novel time series cross-section dataset of personnel transitional justice events in post-authoritarian and post-conflict states. The dataset acknowledges and uses variation of lustration-related activity across time within each country. In doing so, we capture the rich variation in implementation of transitional justice broadly and lustration specifically: whereas 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 East European countries). Furthermore, once introduced, transitional justice legislation seems to prevail in certain countries (as in the Czech Republic and for a long time in the Philippines), whereas in others, it is accompanied by reversals (as in Indonesia).[17] As discussed at length in section 3, our approach leverages variation in the *severity* of lustration, as a way to provide a more valid measure of the phenomenon in question.

[16]Indeed, most contributions to the lustration literature suggest that this transitional justice policy is almost exclusively restricted to Post Communist Europe . This is, however, an artifact of restricting the search for personnel transitional justice to the term 'lustration' and its derivatives. 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 the general "vetting," "purging," or "house-cleaning" to specific terms, such as "de-nazification,""de-communization," or "de-ba'athification." See Mayer-Rieckh and De Greiff 2007; Ellis 1996; Letki 2002; Volčič and Simić 2013; Stan 2013; Stan and Nedelsky 2015.

[17]To see this most clearly, we refer to the readers to Figure 6 discussed in section 3.

2 The Formal Model

Can lustration improve the quality of representation by eliminating the effectiveness of blackmail with dark authoritarian secrets? Our model of incomplete information reconstructs the mechanism through which blackmail interferes with politicians' ability to faithfully represent policy preferences of the electorate. Solving it allows us to examine the extent to which transitional justice indeed reduces the dependence of politicians on their former leading officers.

Formal modeling has scarcely been used to address questions involving transitional justice [18] and, to the best of our knowledge, has not been used for predicting the effect of personnel transitional justice mechanisms on the long term quality of democracy.

Our model is set following a transition to democracy. It features a former collaborator, who is elected to public office (the Politician), and his former Officer, who demands a policy concession and threatens to reveal embarrassing information against the Politician unless the latter provides that concession. This is what we understand as "blackmail."

We will assume that barring blackmail, politicians would implement the policy preferences of their voters. By making this assumption, we are not claiming that politicians always follow voter preferences; however, this assumption is instrumental to isolating the effect of lustration that we seek to identify. The quality of representation suffers when the

[18] A few exceptions are worth mentioning. First, (Ritter and Wolford 2012) use a formal model to analyze bargaining between criminals and international courts. (Nalepa 2010) presents a model of the dynamics between members of the authoritarian regime and opposition, focusing on explaining the former's incentives to step down even when they might be held accountable and the latter's commitments to amnesty. (Dragu 2014) models the effectiveness of democratic versus authoritarian governments in their use of repression. Finally, (Tyson 2018) models the incentives of rank and file members of the security apparatus to defect from obeying order when the autocrat's downfall is imminent.

politician is vulnerable to extortion and implements the policy preferred by his former secret police officer instead of the policy preferred by voters. Lustration is supposed to make blackmail more difficult if not impossible.

The key tension in our model comes from the fact that while the Officer knows whether he possesses the evidence he threatens to release, the Politician does not. Note that the more stringent the lustration law is, the less likely it is that the Officer possesses such incriminating evidence. Therefore, the probability that evidence of the Politician's collaboration remains in the possession of the officer represents the severity of lustration. Restating this in terms of 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.

The model we propose uncovers circumstances under which lustration improves the ability of politicians to represent their constituents. Previewing our results, our model finds that ideological polarization between the blackmailing officer and the politician decreases the quality of representation. This is intuitive, as it indicates that politicians give in to blackmail more frequently when they are asked to deviate more.[19] The model indicates that the quality of representation increases with the severity of lustration and that high quality of representation is easier to achieve as the price of having skeletons in the closet revealed decreases. Thus, our model shows that, ironically, the greater the public's emphasis on former collaborators not being in politics the more influence authoritarian elites have via this blackmail mechanism.

2.1 Sequence of play

The game starts with a move of Nature, which determines with probability $\pi \in (0, 1)$ that the officer has evidence against the politician, and with probability $1 - \pi$ that he does not.

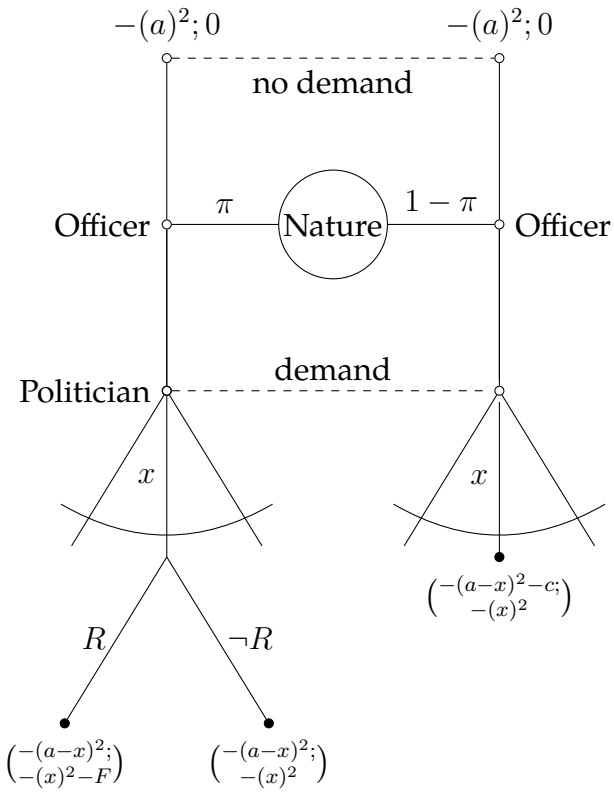
[19]We note, however, that this result is not robust with respect to the functional form of players utility functions (see formal appendix).

π in this model represents the exogenously given severity of lustration, as the more severe lustration is, the less likely it is that unearthed evidence of collaboration indeed remains in the hands of the officer. The *Politician*, P, knows the value of π, but not its specific realization. In the second stage of the game, the *Officer*, (O) decides whether or not to make a policy demand towards the Politician. We represent this decision by one of two actions {demand, no demand}. If no demand is made, the game ends and the Politician implements the policy corresponding to his ideal point.

For simplicity, we assume that the Politician's ideal point is given by $p = 0$ and the Officer's ideal point is represented by $a > 0$. Thus, a captures the ideological distance between the politician and the officer. P observes whether or not a demand has been made, but not whether evidence against him exists (the realization of π). In the third stage, P decides whether to make a policy concession in response to the Officer's demand, and how big this concession should be. We model this concession as proposing a policy $x \in [0, a]$. If evidence against the Politician does not exist, the game ends and the Officer pays the cost of bluffing, $c > 0$. The bluffing cost, c, is the price associated with the officer's pretending that he has evidence which is not actually in his possession. This may be easier to do in a world where media scrutiny is low and merely claiming that a Politician was a collaborator does not require credible evidence to become news. But in a world with free media, a story exposing such fake news would be quickly exposed for what it is. Consequently, one can think of the cost of bluffing as easier in an era where fake news is able to get traction and pass for truth longer before it gets exposed as fake. The cost to the Officer authoring such fake news is that no further attempt of blackmail would be credible.

If evidence does exist, the Officer decides whether or not to reveal the evidence (R or $\neg R$). The revelation is interpreted as the Officer exercising his threat to reveal skeletons in the Politician's closet. In the event that evidence is revealed, the Politician pays the cost of having compromising evidence revealed (which can result in getting fired), F.

Figure 1: Lustration Blackmail Game



2.2 Preferences

The payoffs are a quadratic function of the Euclidean distance between the players' respective ideal points and implemented policy as well as the two types of costs characterized above: (1) the cost to the politician of being fired as a result of revealing skeletons in his closet; and (2) the cost of bluffing incurred by the officer if he makes an empty threat. The utilities, along with the entire game tree, are presented in Figure 1 above.

2.3 Strategies and beliefs

The strategy set of the Politician can be defined as $S_P = \{x(d) : D \rightarrow [0, a]\}$. In other words, a strategy for the politician is a proposal $x \in [0, a]$ that P makes in the event that O

makes a policy demand. Note that O can simply ignore the demand if he choses $x = 0$. [20] While a strategy for the officer can be written as $\sigma_O : T \to D \times \{r(x)\}$, where

$T = \{Evidence, \neg Evidence\}, D = \{demand, \neg demand\}$ and $r(x) \in \{Reject, \neg Reject\}$ $r(x)$ denotes the action taken by O following P's counteroffer x. This part of O's strategy can best be represented as a rejection region $R = \{x : x \in R \to O \text{ reveals}\}$.

This structure resembles that of a signaling game in which O is the sender of the message while P is the receiver. In line with these types of games, the politician may have a chance to update his prior belief about the type of officer he is facing—whether it is an officer equipped with evidence of collaboration or not—after observing the officer's action. Depending on the relative magnitudes of a, c and F, this might be a signaling game or a cheap talk game. The solution concept we use is that of Perfect Bayesian Equilibrium (PBE). This means that strategies are best responses to each other, that we have to specify the *Politician*'s beliefs, and that those beliefs must be consistent with what is observed in equilibrium play.

A possible criticism of how well our model matches the empirical interpretation of blackmail with secret police files is that surely, if someone never consciously collaborated with the secret police, (s)he could be certain that no evidence of his or her collaboration exists. Consequently, "bluffing with secret police files" would be limited to politicians who collaborated with the secret police, while those who are innocent of such collaboration would not be affected. 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.[21] Not infrequently were informal conversations recorded as "reports." The victim of such falsified evidence could therefore

[20] Since P does not observe whether evidence exists or not, but only observes whether a demand was placed or not, we only need a single action do describe his strategy.

[21] Horne 2009; David 2003.

never be certain if evidence of the "collaboration" exists or not.[22] Thus, because of such instances of unintentional collaboration, we assume that theoretically, bluffing is always possible.

2.4 Results and Discussion

We solve this model for separating, pooling, and semi-separating equilibria in the Formal Appendix. Proposition 1 summarizes these equilibria and Figure 2 presents them as a function of F, the Politician's cost of having skeletons in the closet revealed. The set of conditions for the pooling equilibria is distinct from the set of conditions defining the separating and semi-pooling equilibria, with the critical element being the magnitude of F relative to the cost of bluffing, c, the polarization between the Officer and the Politician and the probability that evidence against the politician exists, π.

Proposition 1

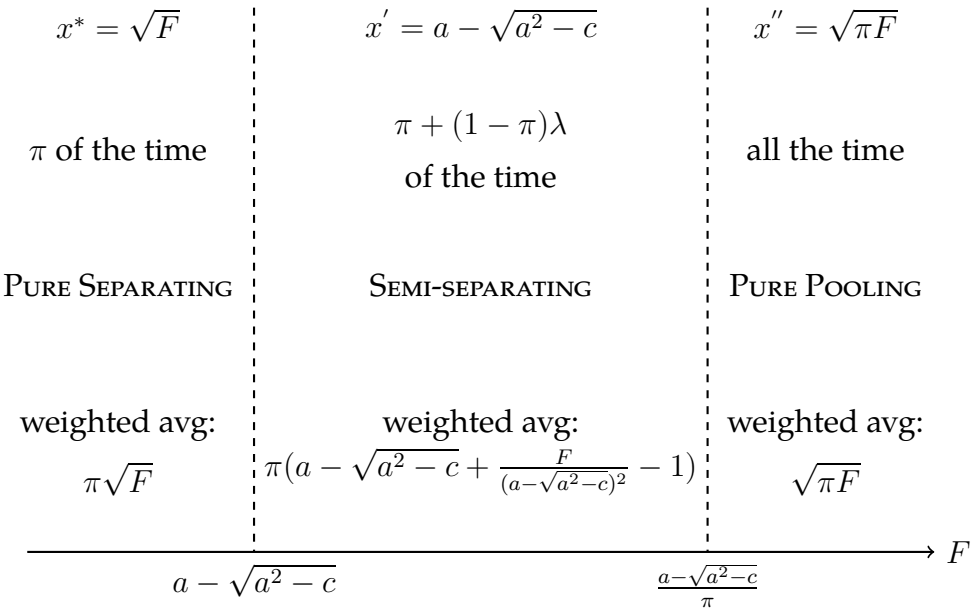
1. Suppose $F < a - \sqrt{a^2 - c}$. There is a pure separating Perfect Bayesian Equilibrium in which the Officer makes a demand if and only if evidence exists and in response, the Politician makes a counteroffer $x^* \equiv \sqrt{F}$. The officer's rejection region in this equilibrium is defined by $R^* \equiv (0, \sqrt{F})$, so he accepts the counteroffer and does not reveal the evidence in his possession. The posterior beliefs of P are described by $Pr(E|\textit{no demand}) = 0, Pr(\neg E|\textit{no demand}) = 1, Pr(E|\textit{demand}) = 1, Pr(\neg E|\textit{no demand}) = 0$.

2. Suppose $F > \frac{a-\sqrt{a^2-c}}{\pi}$. There is a pure pooling Perfect Bayesian Equilibrium in which the Officer always makes a demand, and the Politician responds with a counteroffer $x'' \equiv \sqrt{\pi F}$. The officer's rejection region in this equilibrium is defined by $R'' \equiv (0, \sqrt{\pi F})$ thus he accepts the counteroffer and does not reveal the evidence is his possession. The posterior beliefs of P are described by $Pr(E|\textit{demand}) = \pi, Pr(\neg E|\textit{demand}) = 1 - \pi$.

3. Finally, suppose $a - \sqrt{a^2 - c} \leq F \leq \frac{a-\sqrt{a^2-c}}{\pi}$. There is a semi-separating Perfect Bayesian Equilibrium in which the Officer always places a demand when evidence exists, but if evidence does not exist, he places a demand with probability λ^* and refrains from placing a demand with probability $1 - \lambda^*$. In response to the demand,

[22] Nalepa 2008.

Figure 2: Equilibria



the Politician makes a counteroffer $x' \equiv a - \sqrt{a^2 - c}$, and since the Officer's rejection region is defined by $R' \equiv (0, a - \sqrt{a^2 - c})$, the officer accepts the counteroffer and does not reveal the evidence in his possession. The posterior beliefs of P be described by $Pr(E|\textit{no demand}) = 0, Pr(\neg E|\textit{no demand}) = \lambda^*, Pr(E|\textit{demand}) = 1, Pr(\neg E|\textit{demand}) = 1 - \lambda^*$, where $\lambda^* \equiv \frac{\pi(F-(a-\sqrt{a^2-c}))}{(1-\pi)c}$.

Informally, Proposition 1 states first that for very low costs of revealing skeletons in the closet relative to polarization and the costs of bluffing, there is a separating equilibrium in which the officer only makes a demand if evidence exists and the politician makes a counteroffer $x^* = \sqrt{F}$. Second, it states that if the costs of revealing skeletons are very high relative to the costs of bluffing, there is a pooling equilibrium in which the officer always makes a demand, irrespective of whether evidence exists or not, and the politician makes a counteroffer $x'' = \sqrt{\pi F}$. Third, for moderate costs of revealing skeletons in the closet, there is a semi-pooling equilibrium where the officer sometimes bluffs (i.e., makes a demand even when he has no evidence), and the politician responds with a counteroffer $x' = a - \sqrt{a^2 - c}$. In all equilibria, the counter-offer, when made, is accepted.

For the given set of parameters, all the equilibria are unique provided two plausible

refinements. The first is that neither type play a weakly dominated strategy.[23] The second refinement is that rejection regions that would not be credible for the Officer to play cannot be part of an equilibrium profile.[24]

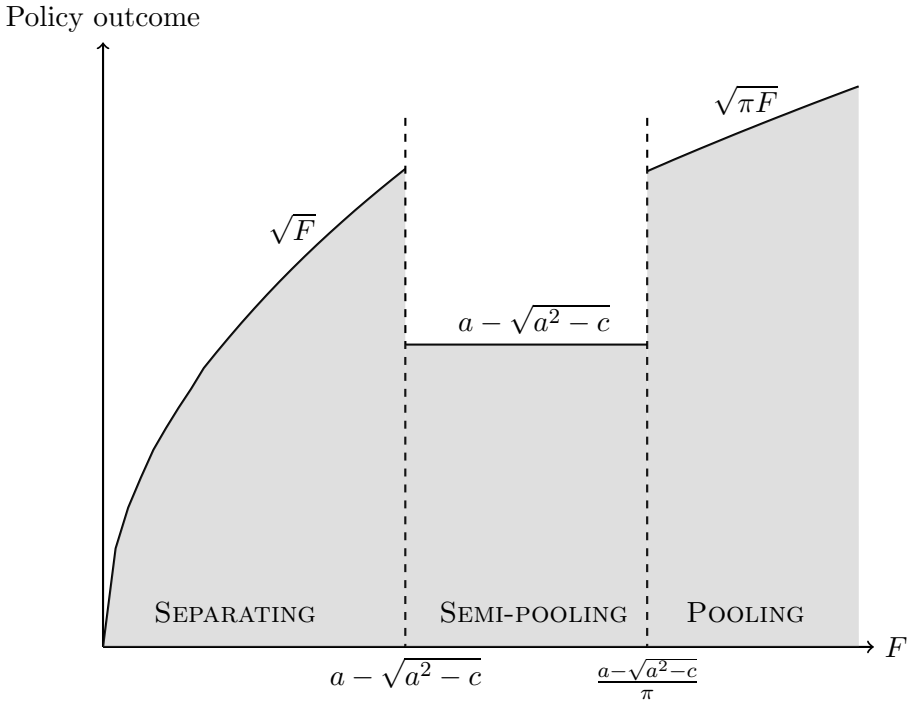
The first implication from this set of propositions is that the quality of representation decreases with a, the distance between the Officer and the Politician. Intuitively, the greater the preference divergence between the two, the more the Politician has to concede to keep the dangerous skeletons from coming out.[25] Second, it is clear that the quality

[23]This refinement ensures that the pooling equilibrium where the Officer regardless of type never makes a demand, sets the rejection region to only include his ideal point, while the politician offers no concessions—is eliminated. In this pooling equilibrium, although we could show that the type with evidence has no profitable deviation, playing no demand when evidence is available is weakly dominated by making a demand.

[24]This refinement eliminates the separating equilibrium in which the type with evidence makes a demand, but the type without does not, the rejection region is set to include only the officer's ideal point, and the politician makes no concession. Note, that without the refinement this would be an equilibrium because the type without evidence cannot improve his payoff, but only lower it by switching to making a demand. The type with evidence cannot improve his payoff and switching actions would involve playing a weakly dominated strategy. The Politician, on the other hand, is getting his highest payoff, so he also has no incentive to deviate. The equilibrium does not satisfy the refinement of not using incredible rejection regions, because a rejection region of a, although it provides the same payoff to the Officer as a rejection region of $[0, \sqrt{F})$ when the Politician makes no concession, is worse than rejection region of $[0, \sqrt{F})$ were the Politician to play anything from the interval $[\sqrt{F}, a]$. Note, that $F < a - \sqrt{a^2 - c}$ is an important restriction here as it prevents the type without evidence from mimicking the type with evidence.

[25]As a robustness check, in the Formal Appendix, we solve the same model but using liner as a opposed to quadratic loss functions to model preferences. The results about dis-

Figure 3: Misrepresentation as a function of cost* of revealing skeletons, F



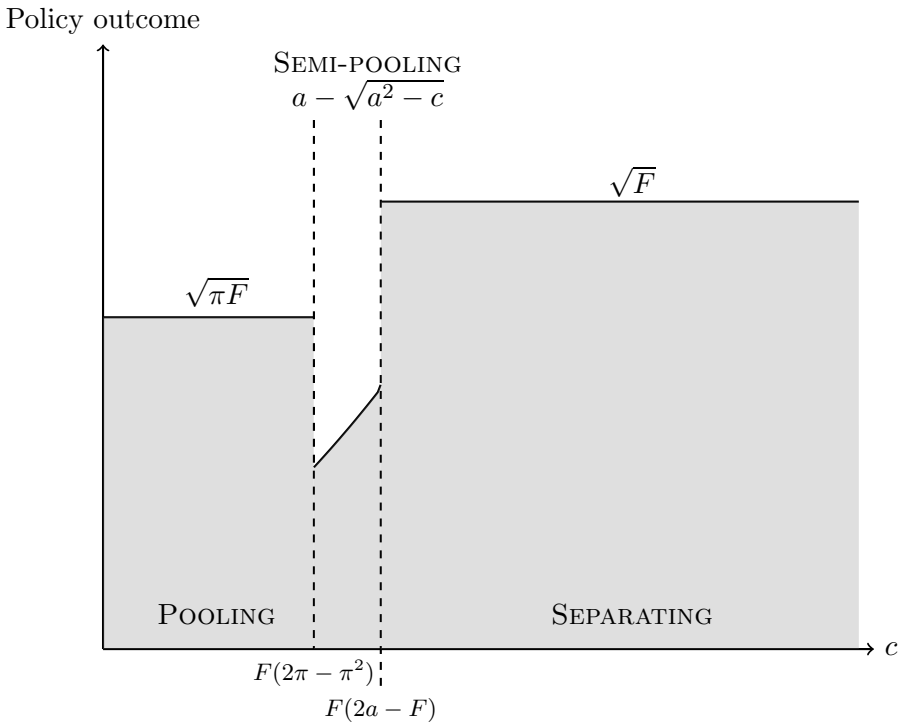
* Assuming $a = 0.7$, $c = 0.4$, and $\pi = 0.55$.

of representation is decreasing in π. This is intuitive because a higher probability that evidence does not exist induces the politician to make a counteroffer closer to his ideal point. Third, quality of representation is highest under the pure separating equilibrium and lowest under the pure pooling equilibrium, indicating that it decreases with the cost of firing. Similarly, a decrease in c, the cost of bluffing, leads to more misrepresentation.

These results are intuitively what we would expect from a model of blackmail. However, the calculation of the separating, pooling, and semi-separating equilibria gives us additional insights allowing us to evaluate the forward-looking argument for lustration.

tance does not hold for this model, but all of the remaining ones do. This extension also allows us to present an "identity result" showing that the average degree of misrepresentation is expressed by the same formula regardless of the equilibrium being played.

Figure 4: Misrepresentation* as a function of cost of bluffing, c



* Assuming $a = 0.7$, $F = 0.35$, and $\pi = 0.55$.
** Note that the figure presents misrepresentation in response to the officer's demand, not total misrepresentation. This means, that were the misrepresentation to be weighed by the frequency of its occurrence, it would be lower under the separating equilibrium (where demands are only made π of the time) than under the pooling equilibrium (where demands are made all of the time).

First, note that lustration works as it "ought to" only under the pure separating equilibrium, where it prevents blackmail entirely: only officers with evidence will succeed at making effective demands. This is because under the pure separating equilibrium, bluffing never occurs and the amount of misrepresentation tracks the amount of evidence left in the officer's possession (so it is directly responsive to the severity of lustration). Both the semi-separating equilibrium and especially the pooling equilibrium, however, allow for blackmail to take place, *even after lustration has been implemented*. The conditions for the pooling equilibrium vis a vis the semi-pooling equilibrium and the semi-pooling equilibrium vis a vis the separating equilibrium are easier to satisfy as the cost of bluffing c decreases, because for higher values of c, it is easier for F to exceed $a - \sqrt{a^2 - c}$ and $\frac{a-\sqrt{a^2-c}}{\pi} < F$, respectively.

Figures 2 and 5 capture these intuitions graphically by calculating for all three equilibria the expected amount of misrepresentation.

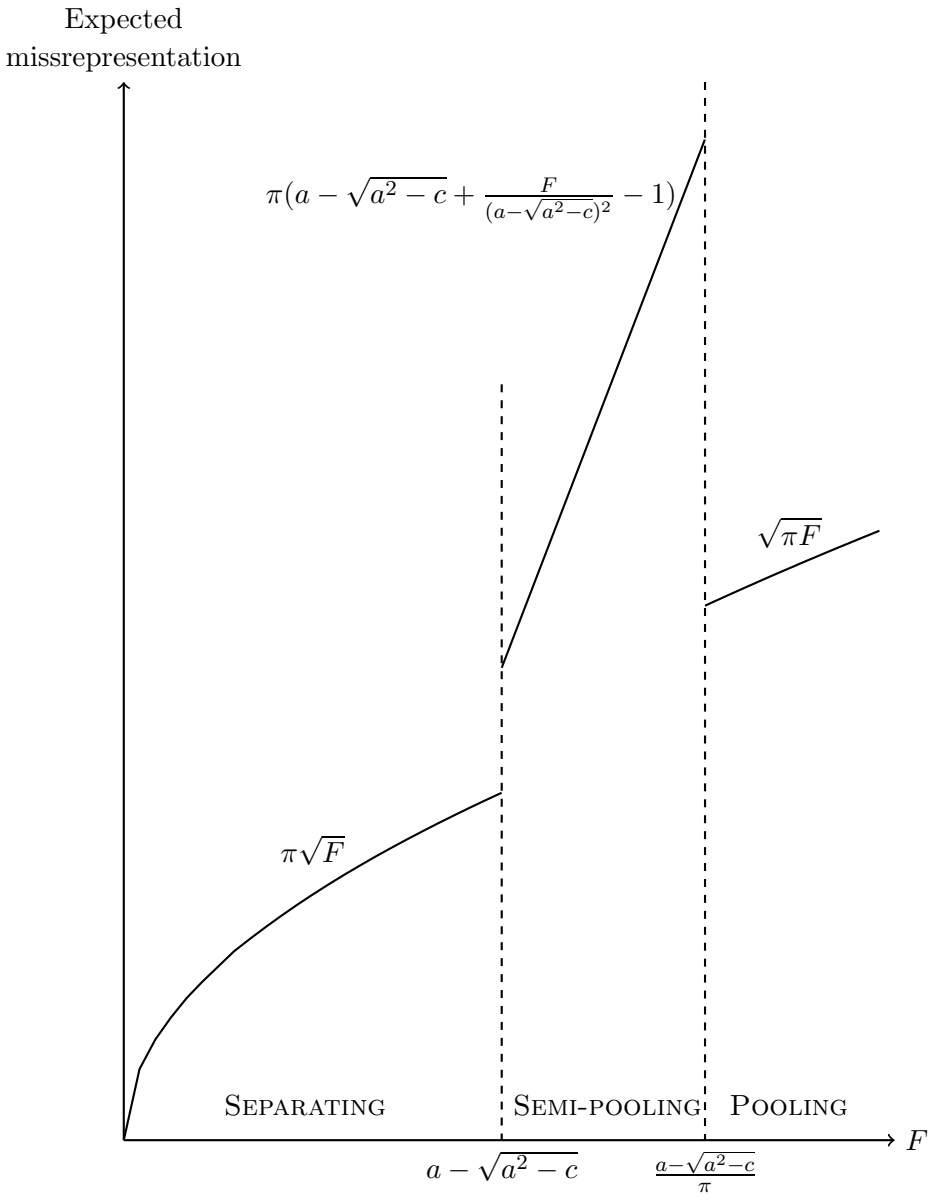
As explained above, in the separating equilibrium, blackmail is restricted to cases where lustration does not reach and leaves evidence in the hands of secret police officers. The more severe lustration, the less blackmail. π of the time, the misrepresentation is $\sqrt{F}$. The expected misrepresentation is then $\pi\sqrt{F}$. But 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 former officers. The frequency of blackmail is completely orthogonal to the actual existence of evidence of collaboration. Under this equilibrium, the extent to which lustration has successfully eliminated evidence of collaboration has no bearing on the effectiveness of blackmail. The expected misrepresentation here is $\sqrt{\pi F}$. Finally, in the semi-pooling equilibrium, the Officer always makes a demand when the evidence exists and with probability λ^* when evidence does not exist. The key feature of the semi-pooling equilibrium is that when it is played, the officer 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. We can calculate the expected departure from perfect representation in this equilibrium. It will be given by: $\pi(a-\sqrt{a^2-c})+(1-\pi)\frac{\pi(F-(a-\sqrt{a^2-c})^2)}{(1-\pi)(a-\sqrt{a^2-c})^2}$, which simplifies to $\pi(a-\sqrt{a^2-c}+\frac{F}{(a-\sqrt{a^2-c})^2}-1)$.

Figure 5 summarizes this argument graphically for $\pi = .65$, $a = 0.55$ and $c = 0.3$.

In this section we have identified conditions under which lustration works as "it ought to" according to the forward looking argument, preventing entirely blackmail with secret police files. Nevertheless, we have also identified conditions —for the pooling and semi-separating equilibria—where lustration is compatible with bluffing and so does not prevent departures from democratic representation. Nevertheless, under all three equilibria, the total expected amount of misrepresentation decreases with lustration. Even when the mechanism through which lustration affects the quality of representation is somewhat different from that which proponents of the forward looking argument maintain, lustration still decreases departures from the politician's ideal point towards the blackmailing officer. Indeed, a somewhat counterintuitive result is that, in terms of F, the lowest level of policy concessions occurs in the pooling equilibrium. Recall, this is the instance where lustration works nothing like the proponents of the forward looking argument expect it to, as blackmail occurs *regardless* of whether evidence exists or not. The intuition behind this intriguing result is as follows: because the politician knows that the chances that evidence against him existing are merely π, he is not willing to put up with the kind of high concessions an officer with evidence would demand (such as the demand in the separating equilibrium, where there is 100% certainty that insubordination would result in skeletons' revelation). The fascinating thing about this equilibrium is that what keeps the secret police officer from extracting more than $\sqrt{\pi F}$ is that the Politician is willing to risk the exposure of his skeletons in the closet rather than put up with an excessively high concession. And of course, after accounting for the frequency with which these concessions occur, it becomes apparent that the separating equilibria, in expectation, lead to the

Figure 5: Expected Misrepresentation* as a function of F



* Assuming $a = 0.55$, $c = 0.3$, and $\pi = 0.65$.

highest quality representation.

The next section derives this as well as additional empirical implications from our model and explains how we illustrate these expectations with a brand new Global Transitional Justice dataset.

3 Empirical Analysis

The game theoretic model presented above represents the dynamic between individual politicians and officers of former authoritarian regimes. The main implication drawn from the model is that lustration prevents blackmail, curbing departures from policies that would be desirable for the electorate. Put differently, we argue that in a perfectly representative democracy, the politician's ideal point would correspond to policies that the voters would like to see implemented. Since the outcome of interest is the distance between the politician's ideal point and the final policy that is implemented, the model predicts individual departures from such perfect representation. Of course, obtaining the necessary evidence to document these departures would require careful analysis of the preferences of individual politicians and former authoritarian elites, a task that would be hard at best, and almost impossible for a large range of countries. Hence, we opt to test the model's implication using a combination of party-level data on ideological positions and quality of representation, and country-level data of lustration policies.

We operationalize our outcome of interest, quality of representation, using a measure that captures both the cohesion within a party, and how salient are the policies that such a party represents. The motivation for this is as follows: The ideological cohesion of a party relies on its members sharing the same ideological platform and supporting policies that advance this platform.[26] If one of these members were blackmailed into making policy concessions to a former agent, the departure from the shared platform would de-

[26]See Kitschelt and Freeze 2010.

crease party cohesion. Hence we can rely on data aggregated at the party level to pick up irregularities in individual behavior.

To measure quality of representation at the party-level, we build on the measure *cosalpo* constructed by Kitschelt et al.[27] The original measure draws from a large-scale survey of experts in 88 different countries, and condenses information on three key components of a total of 506 political parties: cohesion, salience, and polarization.[28] We use a similar measure that leaves out questions pertaining polarization, because it is conceptually close to the distance between the ideal point of the politician and the officer, a key parameter of interest in our model. The result is a score between 0 and 1 that captures the cohesion and salience of each party in the countries of interest. The new proposed measure (*cosal_3*), we argue, accurately captures the *quality of representation* that we refer to in the formal model.[29]

The model's first implication suggests that the probability with which secret evidence against the politician exists (parameter π in the model) decreases the quality of representation because it enables blackmail. Throughout this article, we have contended that lustration publicizes evidence of covert collaboration, so the probability of evidence remaining secret, and susceptible to be used for blackmail, decreases. Therefore, we operationalize the probability of secret evidence of collaboration (π) with country data on lustration policies.

Although most cross-national empirical analyses of transitional justice rely on dichotomous variables to indicate lustration,[30] we argue that this misses variation that is relevant

[27]Kitschelt et al. 2009; Kitschelt and Freeze 2010.

[28]The Empirical Appendix provides full details on how this index was constructed

[29]In the appendix, as a robustness check, we repeated all our analyses using the original cosalpo measure to ensure are findings are not contingent of this newly created measure of ours.

[30]See for example the *Transitional Justice Database Project*, or the *Post-conflict Justice and*

for our theory. Under this approach, countries that implemented a limited lustration policy (for example screening only politicians running for elected top-level positions), would have the same score than countries with a comprehensive lustration policy (a policy that screens all elected positions). Yet it is more likely that secret information would be left undisturbed and ready to be used for blackmail under a limited lustration, rather than under a comprehensive one. Therefore, an adequate measure should reflect the variation on lustration policies across countries.

To capture this variation, we built our own dataset of Global Personnel Transitional Justice. We started with building chronologies of personnel transitional justice events. For each country, we recorded each "nudge"[31] in the direction of promoting or restricting lustration, and coded such events as progressive or regressive, respectively. Progressive events introduce new legislation or broaden the existing one, whereas regressive events attempt to constrain lustration through measures such as voting down or vetoing lustration provisions by the constitutional court. Submitting a lustration proposal to the floor of the legislature, passing or upholding lustration legislation, or overturning a presidential veto against such legislation are all examples of progressive events, whereas amendments decreasing the set of offices subject to lustration are coded as regressive events.[32] The events we analyze here are only those that deal with uncovering *secret* collaboration with the authoritarian regime. Although this is a narrower definition of lustration than that used by other scholars,[33] it is an excellent empirical match for our model, which entails revealing information that nobody except the Officer knows about.

Figure 6 summarizes these data, and shows that indeed there is empirical variation in the lustration policies implemented by each country. Each panel in the figure depicts lus-

Sustainable Peace project.

[31]Thaler and Sunstein 2003.

[32]See the Appendix for a detailed description of the data gathering process.

[33]For example, in the works of Olsen 2010; Horne 2009.

tration events in a given country after transition. The horizontal axis indicates the calendar year and the vertical axis indicates the frequency of events in that year: red, upward lines mark progressive events, and blue, downward lines mark regressive events. Taken together, the panels show that in many countries, lustration was not implemented once and for all. Rather, among lustration bills that were proposed, some were successfully passed through the legislative process, whereas others failed or were eventually overturned (by an executive, an upper chamber, or a constitutional court). Some survived one term only to be struck down by a subsequent government. Figure 6 also suggests that lustration is popular beyond Post-Communist Europe contrary to what the concentration of studies of lustration in Post-Communist Europe may suggest.

In order to maximize the richness of the lustration trajectories presented here, we propose a score that captures the *severity* of lustration. This score is defined as the total number of progressive lustration events divided by the total number of lustration events plus one,[34] taking the value of 0 when country k did not move lustration forward at all, and approaching 1 when the number of progressive events is closer to the total number of events.[35] We add 1 to the denominator to ensure that this score is defined for countries that have had no lustration events. Importantly, this *severity* variable is observed at the country level, whereas the dependent variable, *quality of representation*, is measured at the party level.

We believe that this events-based *severity* score is an improvement over existing measurement, although it is far from ideal. First, our measure of *severity* is time invariant, which means that it cannot explain variation in the quality of democracy over time.[36] Sec-

[34] $severity_k = \frac{progressive\ events_k}{total\ events_k + 1}$

[35] In any given country, the number of regressive transitional events is strictly lower than the number of progressive transitional events. Thus, countries with at least one progressive event will have a non-zero score.

[36] Note, however, that our measure gages how frequently lustration progresses forward

ond, our measure could improve if it accounted for how many past and current positions will be subject to vetting, how severe are the sanctions for collaboration or how many individuals were found exonerated given the numbers investigated. Such detailed information, although clearly superior, would require meticulous data on policy implementation. While collecting it might be feasible for a handful of cases, it is prohibitively costly for the entire sample that we present here. Given our resources, we opted to focus on the theoretical drivers of politicians' behavior: in our model, the Politician's behavior is driven by her beliefs about how likely blackmail is to occur. The higher the proportion of progressive lustration events, as reflected in our measure of *severity*, the more likely it is that they will not be blackmailed with secret police files. With this in mind, our first hypothesis can be stated as follows:

Hypothesis 1 All else kept constant, the quality of representation will increase as the *severity* of lustration increases.

A second empirical implication of the model is that the quality of representation decreases as the cost of having compromising information revealed increases (parameter F in the model). We operationalize this as the electoral costs associated with revealing skeletons in the closet. We argue that in democracies recovering from authoritarian rule, voters care about keeping former secret collaborators out of office, but that the saliency of this issue diminishes over time[37] when holding lustration constant. This claim is supported by literature on historical memory,[38] but is also associated with the fact that in a young

(progressive events) relative to all other actions that could hinder it (regressive events). As such, it captures certain features of lustration that a dynamic measure would miss.

[37]Stan 2006; Bernhard and Kubik 2014; Wilde 1999; Cohen 1995; De Brito, Enríquez, and Aguilar 2001.

[38]Bernhard and Kubik 2014.

democracy, more voters will have experienced (and remember) life under autocracy[39] than in a seasoned democracy. Hence, we operationalize F as the years lapsed since transition,[40] and we propose that the longer the democratic experience of a country, the less attentive will voters be to the fact that their politicians collaborated with the authoritarian regime. When voters pay less attention to collaboration, publicizing such "skeletons" becomes less costly, and politicians will give into blackmail less frequently, regardless of the lustration policies implemented in that country. Formally:

Hypothesis 2 Holding everything else constant, as the time lapsed since the transition to democracy increases, the quality of representation will increase.

It is important to note that this operationalization is somewhat coarse, because the cost of revealing collaboration is not only a function of the how voters react to such allegations, but also of other institutional features. For example, it could be that revelations about collaboration are costlier in electoral systems with the personal vote (such as SMD or OLPR) than in systems where voters cannot directly specify which candidate gets elected into office.[41] Or it could be that voters in some countries are more interested in elections than in other countries, meaning that revelations would be mediated by an overall interest in electoral participation or turnout. Nevertheless, we believe that the political cost of revealing skeletons in a politician's closet should be higher, on average, in recently transitioned countries than countries that transitioned a long time ago.

The third implication of our theory is that the ideological distance between Politicians and Officers negatively impacts the quality of democracy. Similar to the dependent variable, *quality of representation*, our operationalization of distance relies on the DALP survey and it is also measured at the party-level. Specifically, we take advantage of the fact that

[39]Pop-Eleches 2007; Pop-Eleches and Tucker 2011; Pop-Eleches and Tucker 2012.

[40]The date of transition is taken from the Authoritarian Regimes Database (Geddes, Wright, and Frantz 2014), though we have updated it to 2017.

[41]Carey and Shugart 1995.

DALP reports average ideological placements of political parties in countries from our sample. Furthermore, since some of these parties are successor authoritarian parties, we use their placement to proxy the ideological placement of Officers. Therefore, for each non-successor party p in country k, we define *party distance* as the absolute difference in left-right placement between party p and the placement of the country's successor authoritarian party in country k. Our third hypothesis can then be stated as:

Hypothesis 3 As the distance between a country's successor authoritarian party and non-successor parties increases, the quality of representation should decrease.

It is important to note that some countries in our sample do not have a party that is a successor to the authoritarian regime. For countries that had no successor, we replaced the missing distance with the average distance between successor and party across all remaining countries and we included an indicator variable to account for these cases.[42]

Finally, we operationalize parameter c, the cost of bluffing, with the freedom of the press index as reported by Reporters Without Borders.[43] Our logic for this operationalization is as follows: when an officer of the authoritarian regime tries to bluff a politician into thinking that he is in possession of evidence against the politician, such a bluff can be easily exposed in an environment with free media. Put differently, claims regarding the collaboration between a politician and the former secret police will likely be reported, and if the press is free, claims will be eventually investigated and denounced if found to be false. A wide literature discusses how thorough journalists are when it comes to revealing evidence of former secret police conduct.[44] Although the scrutiny offered by journalists is

[42]See Empirical Appendix for replication of our analysis limited to the subsample of countries for which we do have information on the successor party.

[43]To avoid confusion, we linearly transformed this variable to make lower rankings correspond to lower freedom of the press in a given country.

[44]Laplante and Phenicie 2009; Pinto 2010; Chapman 2009.

not perfect, the higher the degree of media independence, the more likely it is that a Officer's bluff will be called.

Our model predicts that an increase in the costs of bluffing increases the chances of an equilibrium with no bluffing, that is, an equilibrium where only politicians with "skeletons in their closets" get blackmailed. Therefore, holding the *severity* of lustration constant, an increase in media freedom should increase the quality of representation because it decreases the amount of bluffing (or eliminates it altogether). Put differently, our model implies that severe lustration directly improves the quality of representation by decreasing the pool of politicians that could be subject to blackmail or bluffing (hypothesis 1), but even for two countries with equal lustration policies, the country with greater media freedom will have the added positive effect on the quality of representation by eliminating bluffing. Thus, we hypothesize:

Hypothesis 4 Holding other tings constant, higher freedom of the media should be associated with higher quality of representation.

3.1 Statistical Models

Our hypotheses can be translated to the following linear specification:[45]

$$\begin{aligned} \textit{quality of representation}_{p,k} = & \beta_k + \beta_d * \textit{party distance}_{p,k} + \beta_s * \textit{severity}_k + \\ & \beta_n * \textit{years since transition}_k + \beta_c * \textit{press freedom}_k + \\ & \beta_X \sum X_k + e_{p,k} + e_k \end{aligned}$$

where *quality of representation* is operationalized using the measure of programmaticness (*cosal3*) for party p in country k. In addition to the main variables of interest described above, the model incorporates a series of country-level control variables ($\sum X_k$), as well as a series of country-specific intercepts (β_k).

[45]See Empirical Appendix for a derivation of this specification.

As implied by the equation above, the outcome of interest (*quality of representation*) and one regressor (*party distance*) are party-level attributes, while the remaining variables are measured at the country-level. To account for this structure, we use multilevel regression. Given that parties are nested within countries, estimating this model using ordinary least squares regression (OLS) would risk biased and inefficient results because characteristics of parties within the same country are not independent of one another other.[46] Multilevel regression can be thought of as a compromise between assuming all parties are independent from each other and assuming that all variation across parties can be explained by the country that they belong to. A multilevel approach will structure the variation in the dependent variable according to party-level and country-level characteristics. Specifically, we will estimate the model above with a series of country-level coefficients (β_j) that will share an underlying distribution with a global mean (μ_{β_j}) and standard deviation (σ_{β_j}).[47] This equation, we argue, is the best way of modeling the outcome of interest given the data available.

Before presenting and discussing the results, we address one possible concern: that countries where lustration was successfully implemented also have parties that are better at representing voters. Were this the case, we would risk overestimating the relationship between lustration severity and quality of representation. Such overestimation, however,

[46]Traditional solutions for data with similar structures to ours (such as fixed effects or clustered standard errors) are inapplicable because some countries only have one party that is not successor authoritarian.

[47]Assuming that all parties are independent of one another would require estimating a fully pooled model with one global intercept. Assuming total dependency on country characteristics would mean a traditional fixed-effects approach that would estimate one intercept per country (minus a base category) . 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_j}.

would have been captured by the country-level intercepts. Put differently, if countries that implemented transitional justice were indeed the same ones whose parties offer higher quality of representation, then all parties in that country should be affected. This effect, in turn, would be captured by the country-specific estimates.

The results of our estimations are shown in Table 1. For ease of interpretation, all variables have been linearly transformed to range between 0 and 1. All models presented here include country-level intercepts (estimations for β_j), and an indicator variable controlling for missing successor parties, both of which are omitted from the table out of space considerations. The table reports the average change in *quality of representation* of a non-successor party that is associated with an increase of one unit in the covariate of interest.

The resulting estimations are generally consistent with the implications of our theoretical model. Most importantly for our argument, they show that higher scores of *severity* of lustration are associated, on average, with more programmatic parties. The effect is statistically significant at the 0.05% confidence level, and robust to the inclusion of country-specific intercepts and other relevant covariates. In addition, its average effect is quite substantial: the difference between a country that did not attempt any lustration at all (*severity* score of 0) and a country with the highest lustration *severity* (score of 0.82) is 0.17 points according to the point estimate of Model 4. Since the variable *quality of representation* ranges from 0 to 1, this average effect accounts for one fifth of the possible range of our outcome of interest.

The models reported here are static, meaning that they take a snapshot of a country, and thus cannot produce evidence of changing quality of representation over time. Nevertheless, they have many redeeming features. First, DALP extends to over 80 democracies and exhibits almost perfect overlap with the set of countries included in our database. No competing datasets (CHES or the Manifesto Project) allow us to measure quality of representation for even half of the countries in the Global Transitional Justice Dataset.

We also find a negative and statistically significant effect for the ideological distance

Table 1: Quality of representation and severity of lustration

	Model 1	Model 2	Model 3	Model 4	Model 5
party distance	-0.103^{***}	-0.107^{***}	-0.109^{***}	-0.111^{***}	-0.111^{***}
	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
severity		0.278^{***}	0.255^{***}	0.208^{***}	0.209^{***}
		(0.062)	(0.064)	(0.069)	(0.070)
years since transition			0.150	0.108	0.113
			(0.103)	(0.107)	(0.115)
press freedom				0.370^{*}	0.369^{*}
				(0.215)	(0.217)
opposition status					0.005
					(0.038)
Constant	0.379^{***}	0.278^{***}	0.217^{***}	-0.039	-0.044
	(0.032)	(0.036)	(0.055)	(0.158)	(0.163)
Country intercepts	Y	Y	Y	Y	Y
Missing successor	Y	Y	Y	Y	Y
Observations	313	313	313	307	307
Log Likelihood	177.450	184.327	184.036	179.721	177.384
Akaike Inf. Crit.	-344.900	-356.654	-354.073	-343.442	-336.768
Bayesian Inf. Crit.	-326.169	-334.177	-327.849	-313.627	-303.226

Note: *p<0.1; **p<0.05; ***p<0.01

between each party and its successor (*party distance*). This finding is consistent with hypothesis 4, which states that as the distance between each party and the successor party increases, the quality of representation should decrease. This result is robust to a number of specifications, most notably using only countries with successor parties (see Empirical Appendix).

This result could, however, be driven by preexisting dynamics between the opposition and authoritarian parties. Suppose, for instance, that parties that existed during the authoritarian period share some ideological characteristics that made their members more likely to be blackmailed. We control for this possible confounder using a measure of how active the opposition was in representative state institutions during the authoritarian era. Cheibub et.al.[48] report data on the status of the opposition in authoritarian legislatures [49] In our sample, only two authoritarian regimes had no legislature at all, so we turned their variable *opposition 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.

Models 3-5 include the variable *years since transition*, which captures time since transition to democracy as a proxy for the cost of having skeletons revealed (F). We hypothesized (hypothesis 2) that over time, politicians pay a lower cost for being exposed as collaborators of the former authoritarian regime. As a result, they should be less prone to blackmail and the quality of representation should increase. However, the coefficient associated with *years since transition* is not significant, which means that we find no evidence to support hypothesis 2. We attribute this lack of support to the coarse measure of costs of firing used in our analysis.

The last two models also include the variable *press freedom*, which is how we opera-

[48]Cheibub, Gandhi, and Vreeland 2010.

[49]Specifically, on whether the authoritarian regime had no legislature, a single party-legislature, or a multiparty-legislature.

tionalize parameter c, the cost of bluffing. We find that the variable is positively related to the quality of representation, although this association is significant at the 90% level. We also find that it does not change the estimated effects of the rest of the theoretically relevant regressors, which we take as evidence supporting our hypotheses. Interestingly, the constant term is no longer statistically significant when *press freedom* is included, and we interpret this as a type of omitted variable bias: in models 1-3, the average effect of a free press and time since transition is encapsulated by this constant.

Figure 7 provides a visual representation of the main relationship of interest in this paper: the one between *quality of representation* and *severity* of lustration. The figure shows the *severity* score on the horizontal axis, and the average country effects estimated in model 1 on the vertical axis. Model 1 included only the regressor that was measured at the party level—*party distance*—and estimated a series of country-specific intercepts that capture the average country-level quality of representation. Figure 7 suggests a positive relationship between these two variables. Furthermore, this pattern seems to be much clearer for the observations that fall in the right half of the panel: the countries that implemented more severe lustration policies. This latter point supports our decision to measure transitional justice as a process, rather than as a dichotomous measure.

Secondly, the evidence shows that the severity of transitional justice, specifically lustration policies, is positively related to the quality of representation. This result is promising insofar as it suggests that lustration policies produce, on average, higher quality of representation after transition. However, the theoretical model also shows that lustration by itself does not prevent blackmail of former dissidents turned politicians. The proper functioning of lustration policies requires an accompanying free press to prevent bluffing.

4 Conclusion

In this paper, we investigated how lustration, understood as the policies that reveal information about prior collaboration with secret authoritarian elites, influence the quality of representation. We used formal theory to model the situation in which officers of the former authoritarian regime use secret police files to blackmail new democratic elites. Such blackmail, in our model, is aimed at forcing the democratic elites to implement policies that are amenable to the former authoritarian elites and not to the voters. 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' "closets." Also, the ideological distance between the former autocrats and the former dissidents increases the ability of former autocrats to extract concessions via blackmail.

Our paper makes a significant contribution to the study of democratic representation. It does so by linking transitional justice to secret authoritarian legacies. Specifically, it finds that lustration decreases the amount of potentially embarrassing information that remains hidden and useful for blackmail and so prevents blackmail of current politicians by former authoritarian elites. We test and confirm this implication empirically using an original dataset of all lustration events in former authoritarian regimes.

Our argument has also important implications for normative transitional justice, in that it contributes to the understanding of the relationship between transitional justice, rule of law and the quality of representation in new democracies. Key for this contribution is showing that transitional justice can enhance representation without jeopardizing the rule of law, because effective lustration means revealing skeletons in the closet, not necessarily banning former collaborators from running for office or criminalizing their former behavior. Many normative transitional justice scholars have argued that transitional justice is retroactive as it undermines the legal principle of "no punishment without a crime." These skeptics have argued that lustration cannot possibly offer a legal foundation of new

democratic states, because it criminalizes acts that were not only deemed legal by the former authoritarian regime, but indeed encouraged. Halmai argues that "living well is the best form of revenge."[50] Other scholars have advocated for the so-called "Spanish Model of Transitional Justice".[51] These authors 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..[52] Misuses of lustration, de-communization, and de-ba'athification corroborate this endorsement of the Spanish model.[53]

We note that, 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, damaging 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".[54] Forgiving and forgetting may sabotage the capacity for elected politicians to represent voters, a phenomenon that is hard to pick up on by studies focusing on the immediate aftermath of transition. Thus, even if one agrees with the arguments of Elster and Holmes question-

[50]Halmai, Scheppele, and McAdams 1997.

[51]Most notably Elster 1998; Holmes 1994; Michnik 2007.

[52]Elster 2004.

[53]See Nalepa 2010; Kritz 1995, for examples on these misuses. Of course, one need look no further that Iraq's policy aimed at purging new democratic institutions of former Ba'athists. De-ba'athification prevented 185 members of Saddam Hussein's party, mostly Sunnis, from running for the legislature in 2003. Despite its promise to promote societal reconciliation, it ignited ethnic tensions. France's policy of *épuration*, banning former Vichy collaborators from holding office after WWII, had very similar effects. See Elster 2004; David 2006.

[54]Nalepa 2010.

ing 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. As such, our argument bears similarity to other formal models showing that legislation with strong normative components, such as human rights protections and legal limits on executive behavior, can be self-enforcing.[55]

Our contributions speak to the effects of transparency policies on the long-term quality of representation in new democracies, an area that has sparked relatively little scholarly attention. With a few exceptions,[56] the use of secrets and blackmail to affect policy has not received much scrutiny from political scientists. Yet, in 2017, the term *kompromat* made front page news as it was used to describe Russia's attempts to intervene in the US elections. Generally, *kompromat* (compromising materials) refers to either embarrassing information or evidence of a person's illegal activity, which could damage that person's career or open her up to prosecution should it be revealed. As Keith Darden points out, this kind of data was routinely collected by authoritarian secret services in the Soviet Era and used to control people through blackmail by threatening the "compromised" with the release of damaging information to the public-or worse, to prosecutors.[57] But such embarrassing or damaging information, even if collected by authoritarian security forces, may be put to use long after the authoritarian regime itself has expired.

Our paper shows that lustration does not always prevent blackmail: we also find that holding all else constant, lack of free press can sabotage transitional justice efforts. Allegations of collaboration need to be investigated and verified: in a liberal democracy, this

[55]Dragu and Polborn 2013, present a model of legal limits and human rights protections that shows how such legislation creates a reluctance on part of administrators to take actions that would violate human rights, even against the executive's wishes.

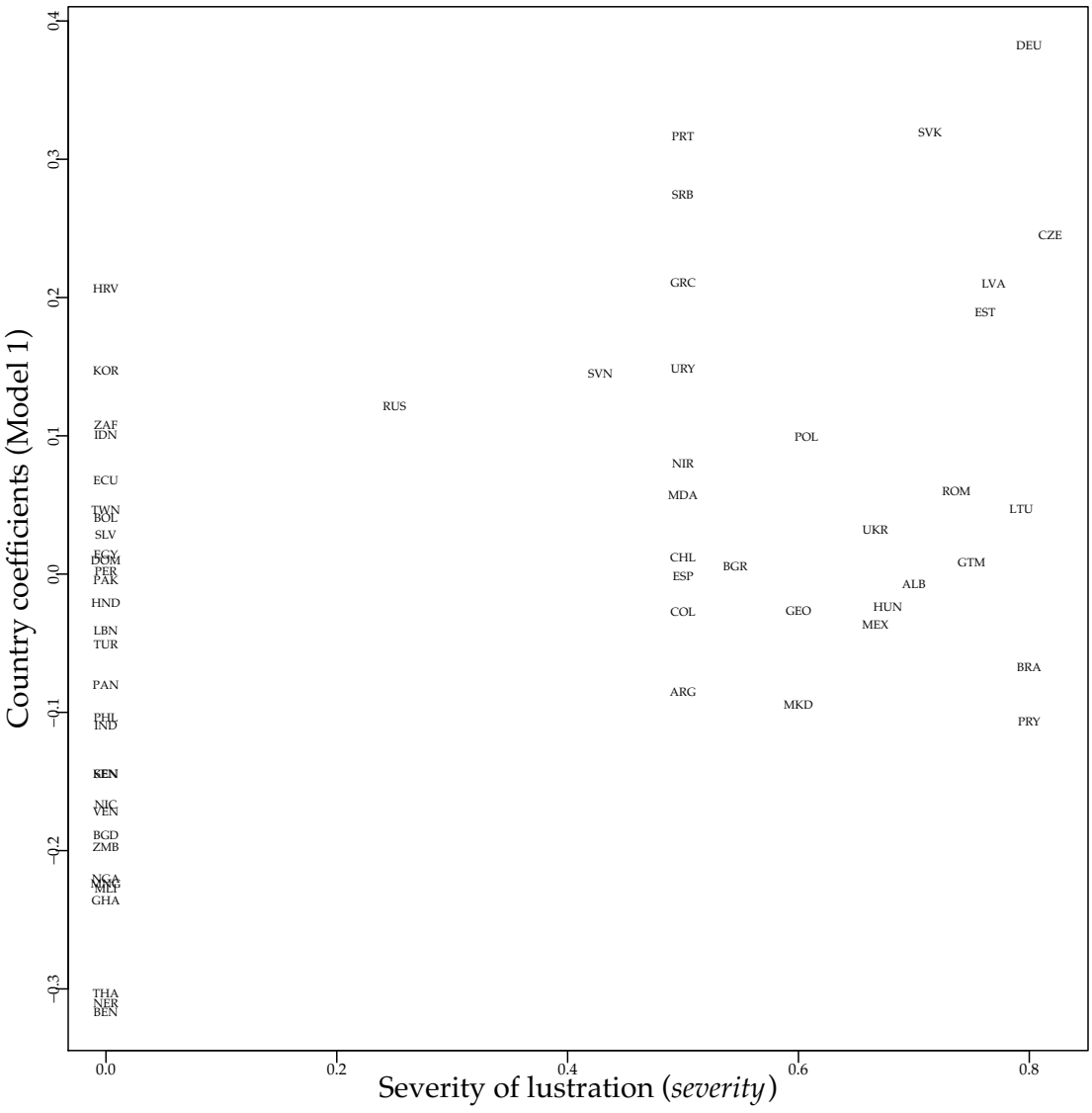
[56]Yarhi-Milo 2013; Felli and Hortala-Vallve 2015.

[57]Darden 2001.

task is often performed by a free press. In the absence of independent media, blackmail occurs through bluffing: threatening to accuse current politicians of having ties with the former authoritarian regime even when no actual evidence exists.

Figure 6: Summary of regressive and progressive lustration events

Figure 7: Severity of lustration and average quality of democracy
Intercept by country



References

Albertus, Michael and Victor Menaldo (2014). "The Political Economy of Autocratic Constitutions". In: *Constitutions in Authoritarian Regimes*, pp. 53–82. DOI: https://doi.org/10.1017/CBO9781107252523.006.

Bernhard, Michael and Jan Kubik (2014). *Twenty Years After Communism: The Politics of Memory and Commemoration*. Oxford University Press.

Carey, John M. and Matthew Soberg Shugart (1995). "Incentives to cultivate a personal vote: A rank ordering of electoral formulas". In: *Electoral studies* 14.4, pp. 417–439. DOI: https://doi.org/10.1016/0261-3794(94)00035-2. (Visited on 04/12/2016).

Carroll, Royce and Monika Nalepa (2014). "Can parties represent after communism? The dissipation of the regime divide cleavage in post-Communist party systems". In: *21st International Conference of Europeanists*. Ces. (Visited on 04/01/2016).

— (2016). "Representative Agents, or Electoral Vehicles The under Institutional Underpinnings of Parties' IdeologicalCohesion". In:

Chapman, Audrey R (2009). "Truth finding in the transitional justice process". In: *Assessing the impact of transitional justice: challenges for empirical research*, pp. 91–114.

Cheibub, José Antonio, Jennifer Gandhi, and James Raymond Vreeland (2010). "Democracy and dictatorship revisited". In: *Public Choice* 143.1-2, pp. 67–101.

Cohen, Stanley (1995). "State crimes of previous regimes: Knowledge, accountability, and the policing of the past". In: *Law & Social Inquiry* 20.1, pp. 7–50. DOI: https://doi-org.proxy.uchicago.edu/10.1111/j.1747-4469.1995.tb00681.x.

Darden, Keith A (2001). "Blackmail as a Tool of State Domination: Ukraine under Kuchma". In: *E. Eur. Const. Rev.* 10, p. 67. URL: https://heinonline.org/HOL/P?h=hein.journals/eeurcr10&i=171.

David, Roman (2003). "Lustration Laws in Action: the Motives and Evaluation of Lustration Policy in the Czech Republic and Poland (1989–2001)". In: *Law & Social Inquiry* 28.2, pp. 387–439. DOI: https://doi-org.proxy.uchicago.edu/10.1111/j.1747-4469.2003.tb00197.x.

— (2006). "From Prague to Baghdad: Lustration Systems and their Political Effects". In: *Government and Opposition* 41.3, pp. 347–372. DOI: https://doi.org/10.1111/j.1477-7053.2006.00183.x. (Visited on 07/16/2015).

— (2011). *Lustration and transitional justice: Personnel systems in the Czech Republic, Hungary, and Poland*. University of Pennsylvania Press.

De Brito, Alexandra Barahona, Carmen González Enríquez, and Paloma Aguilar (2001). *The Politics of Memory and Democratization: Transitional Justice in Democratizing Societies: Transitional Justice in Democratizing Societies*. Oxford University Press. (Visited on 08/21/2014).

Dragu, Tiberiu (2014). "On repression and its effectiveness". In: *Journal of Theoretical Politics* 29, pp. 599–622. DOI: https://doi.org/10.1177/0951629817710563.

Dragu, Tiberiu and Mattias Polborn (2013). "The Administrative Foundation of the Rule of Law". In: *The Journal of Politics* 75.4, pp. 1038–1050. URL: https://www.journals.uchicago.edu/doi/abs/10.1017/S002238161300100X.

Ellis, Mark S (1996). "Purging the past: The current state of lustration laws in the former communist bloc". In: *Law and Contemporary Problems* 59.4, pp. 181–196.

Elster, Jon (1998). "Coming to terms with the past. A framework for the study of justice in the transition to democracy". In: *European Journal of Sociology* 39.01, pp. 7–48.
— (2004). *Closing the books: Transitional justice in historical perspective*. Cambridge University Press. (Visited on 03/29/2016).
Escribà-Folch, Abel and Joseph Wright (2015). "Human rights prosecutions and autocratic survival". In: *International Organization* 69.2, pp. 343–373. DOI: https://doi.org/10.1017/S0020818314000484.
Felli, Leonardo and Rafael Hortala-Vallve (2015). "Collusion, Blackmail and Whistle-blowing". In: *Quarterly Journal of Political Science* 11, pp. 279–312.
Gandhi, Jennifer and Adam Przeworski (2007). "Authoritarian Institutions and the Survival of Autocrats". In: *Comparative Political Studies* 40.11, pp. 1279–1301. DOI: https://doi.org/10.1177/0010414007305817.
Geddes, Barbara, Joseph Wright, and Erica Frantz (2014). "Autocratic Breakdown and Regime Transitions: A New Data Set". In: *Perspectives on Politics* 12.2, pp. 313–331. DOI: https://doi.org/10.1017/S1537592714000851.
Halmai, Gabor, Kim Lane Scheppele, and A. James McAdams (1997). *Transitional Justice and the Rule of Law in New Democracies*. University of Notre Dame Press Notre Dame, IN.
Holmes, Stephen (1994). "End of Decommunization, The". In: *E. Eur. Const. Rev.* 3, p. 33.
Horne, Cynthia M (2009). "International Legal Rulings on Lustration Policies in Central and Eastern Europe: Rule of Law in Historical Context". In: *Law & Social Inquiry* 34.3, pp. 713–744. DOI: https://doi-org.proxy.uchicago.edu/10.1111/j.1747-4469.2009.01162.x.
Horne, Cynthia M. (2017). *Building Trust and Democracy in Transition: Assessing the Impact of Transitional Justice in Post-Communist Countries*. Oxford University Press.
Kaminski, Marek M., Monika Nalepa, and Barry O'Neill (2006). "Normative and strategic aspects of transitional justice". In: *Journal of Conflict Resolution*, pp. 295–302. (Visited on 08/21/2014).
Keefer, Philip (2007). "Clientelism, Credibility, and the Policy Choices of Young Democracies". In: *American Journal of Political Science* 51.4, pp. 804–821. URL: https://www.jstor.org/stable/4620101.
Kitschelt, Herbert and Kent Freeze (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*. Vol. Available at https://goo.gl/xhxeyj.
Kitschelt, Herbert and Steven I. Wilkinson (2007). *Patrons, clients and policies: Patterns of democratic accountability and political competition*. Cambridge University Press. (Visited on 04/11/2016).
Kitschelt, Herbert et al. (2009). "Measuring democratic accountability: An initial report on an emerging data set". In: *Revista de Ciencia Política* 29.3, pp. 741–773. URL: http://www.redalyc.org/html/324/32414667004/ (visited on 04/01/2016).
Kritz, Neil J. (1995). *Transitional justice: how emerging democracies reckon with former regimes*. Vol. 2. US Institute of Peace Press. (Visited on 08/21/2014).
Laplante, Lisa J and Kelly Phenicie (2009). "Mediating post-conflict dialogue: the media's role in transitional justice processes". In: Available at https://goo.gl/tvDDs6.

Lessing, Benjamin and Graham Denyer Willis (forthcoming). "Legitimacy in Criminal Governance: How to Build a Drug Empire From Behind Bars". In: *American Political Science Review*.

Letki, Natalia (2002). "Lustration and Democratisation in East-Central Europe". In: *Europe-Asia Studies* 54.4, pp. 529–552.

Lundy, Patricia and Mark McGovern (2008). "Whose justice? Rethinking transitional justice from the bottom up". In: *Journal of Law and Society* 35.2, pp. 265–292. DOI: https://doi-org.proxy.uchicago.edu/10.1111/j.1467-6478.2008.00438.x. (Visited on 08/21/2014).

Mainwaring, Scott (1999). *Rethinking Party Systems in the Third Wave of Democratization: the Case of Brazil*. Stanford University Press.

Mallinder, Louise (2008). *Amnesty, Human Rights and Political Transitions: Bridging the Peace and Justice Divide*. en. Hart Publishing Limited. ISBN: 978-1-84113-771-1.

Mayer-Rieckh, Alexander and Pablo De Greiff (2007). *Justice as prevention: vetting public employees in transitional societies*. Social Science Research Council.

Medina, Luis Fernando and Susan Stokes (2007). "Monopoly and Monitoring: an Approach to Political Clientelism". In: *Patrons, clients, and policies*, pp. 68–83. URL: https://doi.org/10.1017/CBO9780511585869.003.

Merwe, Hugo Van der, Victoria Baxter, and Audrey R. Chapman (2009). *Assessing the impact of transitional justice: Challenges for empirical research*. US Institute of Peace Press. (Visited on 04/01/2016).

Michnik, Adam (2007). "The Polish witch-hunt". In: *New York Review of Books* 54.11, p. 25. DOI: https://doi.org/10.1007/s11127-009-9491-2.

Nalepa, Monika (2008). "To Punish the Guilty and Protect the Innocent Comparing Truth Revelation Procedures". In: *Journal of Theoretical Politics* 20.2, pp. 221–245. DOI: DOI: 10.1177/0951629807085819.

— (2010). *Skeletons in the closet: Transitional justice in post-communist Europe*. Cambridge University Press. (Visited on 03/29/2016).

Nedelsky, Nadya (2013). "From Velvet Revolution to Velvet Justice: The Case of Slovakia". In: *After Oppression*, pp. 390–417. DOI: https://doi.org/10.18356/abdf8199-en.

Olsen Tricia D., Leigh A. Payne Andrew G. Reiter (2010). "Transitional Justice In Balance: Comparing Processes, Weighing Efficacy". In:

Pettai, Eva-Clarita and Vello Pettai (2014). *Transitional and retrospective justice in the Baltic States*. Cambridge University Press. (Visited on 04/01/2016).

Pinto, António Costa (2010). "The Authoritarian Past and South European Democracies: An Introduction". In: *South European Society and Politics* 15.3, pp. 339–358. DOI: https://doi.org/10.1080/13608746.2010.513598.

Pop-Eleches, Grigore (2007). "Historical legacies and post-communist regime change". In: *Journal of Politics* 69.4, pp. 908–926. URL: https://www.journals.uchicago.edu/doi/abs/10.1111/j.1468-2508.2007.00598.x.

— (2010). "Throwing Out the Bums: Protest Voting and Unorthodox Parties after Communism". In: *World Politics* 62.2, pp. 221–260. DOI: https://doi.org/10.1017/S0043887110000043.

Pop-Eleches, Grigore and Joshua A Tucker (2011). "Communism's shadow: postcommunist legacies, values, and behavior". In: *Comparative Politics* 43.4, pp. 379–408. DOI: https://doi.org/10.5129/001041511796301588.

— (2012). "Associated with the past? Communist legacies and civic participation in postcommunist countries". In: *East European Politics & Societies*, p. 0888325412465087. DOI: https://doi.org/10.1177/0888325412465087.

Ritter, Emily Hencken and Scott Wolford (2012). "Bargaining and the Effectiveness of International Criminal Regimes". In: *Journal of Theoretical Politics* 24.2, pp. 149–171. DOI: https://doi.org/10.1177/0951629811429048.

Stan, Lavinia (2006). "The Vanishing Truth? Politics and Memory in Post-communist Europe". In: *East European Quarterly* 40.4, pp. 383–409. DOI: https://search.proquest.com/docview/195177667?accountid=14657.

— (2012). "Witch-hunt or Moral Rebirth? Romanian Parliamentary Debates on Lustration". In: *East European Politics & Societies* 26.2, pp. 274–295. DOI: https://doi.org/10.1177/0888325411403922.

— (2013). "Reckoning with the Communist past in Romania: A scorecard". In: *Europe-Asia Studies* 65.1, pp. 127–146. DOI: https://doi.org/10.1080/09668136.2012.698052.

Stan, Lavinia and Nadya Nedelsky (2015). *Post-communist Transitional Justice: Lessons from Twenty-five Years of Experience*. Cambridge University Press.

Stan, Lavinia and others (2009). *Transitional justice in Eastern Europe and the former Soviet Union: Reckoning with the communist past*. Routledge. (Visited on 08/21/2014).

Tavits, Margit (2005). "The Development of Stable Party Support: Electoral Dynamics in Post-Communist Europe". In: *American Journal of Political Science* 49.2, pp. 283–298. URL: https://www.jstor.org/stable/3647677.

Thaler, Richard H and Cass R Sunstein (2003). "Libertarian Paternalism". In: *American Economic Review* 93.2, pp. 175–179. DOI: https://doi.org/10.1257/000282803321947001.

Thoms, Oskar NT, James Ron, and Roland Paris (2008). "The effects of transitional justice mechanisms: A summary of empirical research findings and implications for analysts and practitioners". In: *Ottawa: Centre for International Policy Studies* Available at https://goo.gl/EFnUjj.

Tyson, Scott A (2018). "The Agency Problem Underlying Repression". In: *The Journal of Politics* 80.4, pp. 1297–1310. DOI: https://doi.org/10.1086/698887.

Volčič, Zala and Olivera Simić (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, Alexander (1999). "Irruptions of memory: expressive politics in Chile's transition to democracy". In: *Journal of Latin American Studies* 31.02, pp. 473–500. URL: https://www.cambridge.org/core/journals/journal-of-latin-american-studies/article/irruptions-of-memory-expressive-politics-in-chiles-transition-to-democracy/1AFD03BA8793ACB237A57EAFC610B219.

Yarhi-Milo, Keren (2013). "Tying Hands Behind Closed Doors: the Logic and Practice of Secret Reassurance". In: *Security Studies* 22.3, pp. 405–435. DOI: https://doi.org/10.1080/09636412.2013.816126.