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

Appendix A: Equilibrium Analysis of the Formal Model

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.

A.1 Pure separating equilibrium

Since it is never rational for A to not make a demand when he has evidence against the dissident,[1] 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 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, 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.

[1] Note 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.

A.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: $Pr(E|0) = p, Pr(\sim E|0) = (1-p)$[2]. 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 me the case that $EU_D(x|0,0) \geq EU_D(d|0,0)$, which is equivalent to $-|d-x| \geq p(-|d-d|-F) + (1-p) * 0$.[3] 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) equillibrium.

A.2.1 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 $(1, R, \lambda; x')$.[4] Note the beliefs consistent with this semi-pooling equilibrium are: $Pr(E|0) = 0), Pr(\sim E|1) = \lambda), Pr(\sim E|0) = 1, Pr(\sim E|0) = 1 - \lambda$. 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]Note 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

[3]Note 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.

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

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:

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

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

$$d-x' < F < \frac{d-x'}{p} \tag{2}$$

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 1:

$$\lambda = \frac{p(F-c)}{(1-p)c} \tag{3}$$

6. To get scope conditions for the semi-pooling equilibrium, we can substitute x' into condition 2:

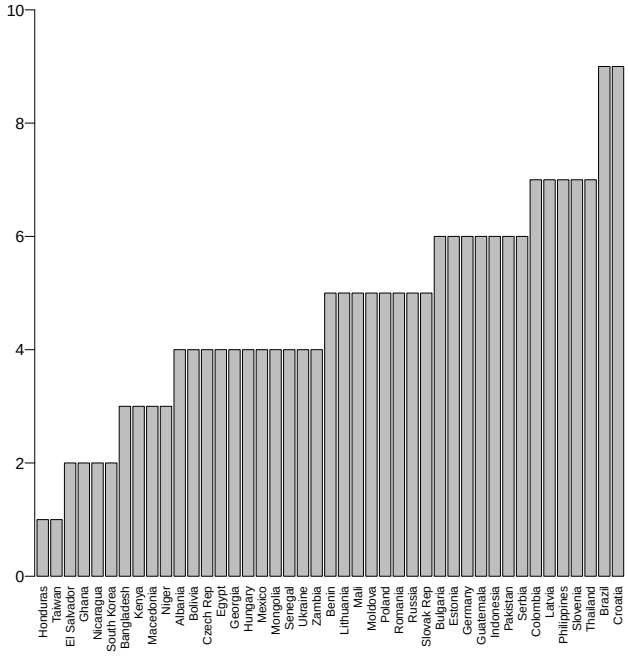
$$c < F < cp$$

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

Appendix B: Data Creation and Summary Statistics

We include a total of 201 cases from 43 countries. The original dataset of transitional justice includes data from Paraguay and Montenegro, but these two cases were excluded from the analysis. Paraguay had a short democratic lifespan and Montenegro's history is highly correlated with Serbia's.

Figure 1: Parties by country included in the analysis



B.1 Programaticness ($cosal\_3$)

We draw from Kitschelt et.al. to operationalize our dependent variable, programaticness. Following their approach, we measure programmaticness with two dimensions: cohesion and salience. For purposes of minimizing endogeneity, we do not include polarization as a relevant dimensions. We use their original raw dataset that reports the survey conducted among experts regarding the characteristics and behavior of specific parties. For both dimensions, we draw from the next set of questions (or issues):

- d1: Party policy position on social spending on a 1-10 scale
- d2: Party policy position on the state role in economy on a 1-10 scale
- d3: Party policy position on public spending on a 1-10 scale
- d4: Party policy position national identity on a 1-10 scale
- d5: Party policy position on traditional authority, institutions, customs on a 1-10 scale

Following Kitschelt et.al. , we measure the cohesion for each issue i of party p in country c by calculating the standard deviation of the responses of that issue. The standard deviation of those issues that received less than 5 scores and of those issues that had a higher score than 3.5 are capped at 3.5 (to avoid outliers resulted from a low response rate). Then, such score was transformed so that higher standard deviation was translated into lower values for cohesion and we normalized them so that they would range between 0 and 1. The salience for issue i of party p in country c is the proportion of experts that gave a valid answer in each of the questions. We then create a $cosal$ measure for each of these issues. Finally, the variable $cosal\_3$

is created by taking the average between three values, $cosal_{d4}$, $cosal_{d5}$, and the maximum value of $cosal_{d1}$, $cosal_{d2}$, and $cosal_{d3}$.

B.2 Severity of transitional justice ($severity$)

We draw from an original dataset of transitional events, which includes transitional justice events for all former party-based authoritarian countries since World War II. The transitional justice vents were gathered using the Keesings Record of World Events, Lexis Nexis Academic Universe, and numerous secondary sources. For each country per year after transition, coders were asked to record all positive and negative transitional events.

A positive transitional event is identified as actions from the government aimed at implementing lustration. These actions can include, for example, submitting a lustration proposal to the floor of the legislature, passing or upholding such legislation, or overturning a presidential veto against it. It can also include widening the scope of an existing law, either by increasing the number of lustrable offenses or the set of people that can be targeted.

A negative transitional event, in contrast, is the limitation, voting down, vetoing or striking down any lustration provision. Attempts to narrow the set of targets or lustrable offenses are also considered negative transitional justice events. Each 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.

We the use such data to create a cross-sectional measure of transitional justice for analysis. For each country, we report the sum of all positive and all negative transitional events, and we create a score of the $severity$ of lustration in each country c:

$severity_c = \frac{total\_positive\_events_c}{total\_events_c+1}$.

B.3 Additional control variables

We also use the next control variables (all the variables are linearly transformed to range between 0 and 1 for easier interpretation).

- d_party: Ideological distance between each party and the successor party of each country., When 2 or more successor parties were present, we calculated the average. When no successor party is present, we inputted the mean distance.
- n_year: Number of years since transition. Own calculations.
- $press\_freedom$: Freedom of the press Index, provided by Reporters without Borders. Data was linearly transformed so that higher values would mean higher press freedom.
- opp_status: Status of opposition during authoritarian regime, data provided by Gandhi, Przeworski, and Vreeland. We transform the original variable into an indicator variable where 0 means that there is no legislature, its members are nonpartisan, or only members from regime party are allowed, and 1 stands for the presence of multiple parties in the legislature during authoritarian period.
- no_succ: No successor. This variable was created using our own coding of successor parties, based on Loxton, Kitschelt, and Loftus.

B.4 Summary statistics and correlation

Table 1: Summary statistics for party-level variables

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	N
$cosal\_3$	0.00	0.21	0.38	0.37	0.51	0.80	201
d_party_3	0.00	0.18	0.27	0.32	0.45	1.00	201

Table 2: Summary statistics for country-level variables

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	N
$severity$	0.00	0.42	0.56	0.50	0.72	0.90	43
n_year	0.12	0.46	0.50	0.55	0.59	1.00	43
$press\_freedom$	0.54	0.71	0.79	0.78	0.85	1.00	43
opp_status	0.00	0.00	0.00	0.49	1.00	1.00	43
no_succ	0.00	0.00	0.00	0.23	0.00	1.00	43

Table 3: Correlation between country-level variables

	$severity$	n_year	$press\_freedom$	opp_status	no_succ
$severity$	1				
n_year	0.02	1			
$press\_freedom$	0.24	-0.04	1		
opp_status	-0.17	-0.31	-0.04	1	
no_succ	-0.02	0.17	-0.34	0.01	1

Appendix C: Robustness checks

We present three robustness checks. The first one provides an alternative measure of transitional justice that measures transitional justice as the interaction between positive and negative transitional events. The second and third robustness checks deal with issues raised by the variable d_party.

C.1 Alternative measure of transitional justice

We estimate the same models presented in the paper but using the interaction between positive and negative transitional events.

Table 4: Explaining programaticness (country intercepts not shown)

	Transitional justice as positive*negative events						
	Model 1a	Model 2.1a	Model 2.2a	Model 2.3a	Model 3	Model 4	Model 5
d_party	−0.045	−0.047	−0.047	−0.048	−0.050	−0.054	−0.054
	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)
$positive\_events$		0.256***		0.312*	0.427**	0.358*	0.397*
		(0.089)		(0.182)	(0.192)	(0.195)	(0.204)
$negative\_events$			0.210**	−0.065	0.267	0.194	0.161
			(0.093)	(0.184)	(0.282)	(0.283)	(0.291)
$positive * negative$					−0.589	−0.470	−0.467
					(0.369)	(0.373)	(0.380)
n_year					−0.153	−0.158	−0.114
					(0.127)	(0.125)	(0.137)
$press\_freedom$						0.355	0.309
						(0.249)	(0.266)
opp_status							0.040
							(0.056)
no_succ							−0.037
							(0.067)
Constant	0.363***	0.298***	0.320***	0.298***	0.358***	0.101	0.099
	(0.031)	(0.037)	(0.036)	(0.037)	(0.083)	(0.198)	(0.219)
Observations	201	201	201	201	201	201	201
Log Likelihood	114.694	117.090	115.705	116.373	117.221	117.756	114.382
Akaike Inf. Crit.	−221.389	−224.179	−221.410	−220.746	−218.443	−217.512	−206.764
Bayesian Inf. Crit.	−208.176	−207.663	−204.893	−200.926	−192.016	−187.783	−170.427

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

C.2 Subset of countries with no missing successors

We also estimate the same models presented in the paper but only in the subset of countries for which we have observed a successor:

Table 5: Explaining programaticness (country intercepts not shown)

	Model 1d	Model 2d	Model 3d	Model 4d	Model 5d
d_party	−0.049	−0.049	−0.047	−0.052	−0.053
	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)
$severity$		0.247**	0.289***	0.239**	0.236**
		(0.103)	(0.095)	(0.096)	(0.102)
n_year			−0.454***	−0.441***	−0.449***
			(0.162)	(0.157)	(0.171)
$press\_freedom$				0.487*	0.499*
				(0.283)	(0.302)
opp_status					−0.008
					(0.062)
no_succ	0.386***	0.261***	0.477***	0.105	0.105
	(0.035)	(0.061)	(0.096)	(0.235)	(0.239)
Observations	146	146	146	146	146
Log Likelihood	87.087	88.453	91.189	92.294	90.430
Akaike Inf. Crit.	−166.174	−166.905	−170.378	−170.589	−164.860
Bayesian Inf. Crit.	−154.240	−151.987	−152.476	−149.703	−140.991

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

C.3 Distance normalized by country

In addition, we present the same set of models but transforming the variable of d_party such that parties within each country range between 0 and 1. Given that this is a within-country linear transformation, we do not find any significant changes from the original model presented here.

Table 6: Explaining programaticness (country intercepts not shown)

	d_party normalized by country				
	Model 1e	Model 2e	Model 3e	Model 4e	Model 5e
d_party_normal	−0.004	−0.004	−0.003	−0.004	−0.004
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
$severity$		0.279***	0.296***	0.237**	0.234**
		(0.107)	(0.098)	(0.102)	(0.108)
n_year			−0.447***	−0.459***	−0.466**
			(0.173)	(0.168)	(0.182)
$press\_freedom$				0.484*	0.494
				(0.293)	(0.313)
opp_status					−0.007
					(0.062)
no_succ	0.373***	0.229***	0.452***	0.099	0.100
	(0.035)	(0.064)	(0.104)	(0.237)	(0.241)
Observations	144	144	144	144	144
Log Likelihood	84.438	86.281	88.556	89.578	87.714
Akaike Inf. Crit.	−160.876	−162.562	−165.111	−165.156	−159.428
Bayesian Inf. Crit.	−148.997	−147.713	−147.293	−144.368	−135.670
Note:					*p<0.1; **p<0.05; ***p<0.01