

PRELIMINARY DRAFT  
**AUTHORITARIAN BACKSLIDING**

MONIKA NALEPA

GEORG VANBERG

CATERINA CHIOPRIS

**ABSTRACT.** A prominent contemporary phenomenon is an apparent “backsliding” of democratic countries into (semi-)authoritarian practices. Importantly, such episodes unfold over time, and often involve uncertainty about the ultimate intentions of governments. Governments typically do not attempt to engage in authoritarian practices immediately, but rather initiate policies or institutional reforms that may in the future facilitate or enable actions that are inconsistent with liberal democratic practices. Building on recent work by Milan Svolik (2018), we develop a formal model that explores both features (and their interaction) in a two-period game in which a government takes an action in period 1 that may allow for subsequent actions that are inconsistent with the rule of law in period 2. Citizens face uncertainty over the ultimate intentions of the government, and must decide whether to replace the first-period government before period 2. The model generates several insights. First, consistent with other work (e.g., Svolik 2018), it suggests that polarization among citizens and elites is an important factor in driving authoritarian backsliding. Extending this logic, the model demonstrates that the degree of polarization necessary to generate the potential for backsliding depends critically on the uncertainty facing citizens about the type of incumbent. Finally, in such a setting, citizens may support incumbent governments even if there is some risk that these governments are “closet autocrats” despite the fact that citizens are fundamentally opposed to authoritarianism. One consequence is that in our model, citizens may genuinely come to regret their electoral choices—a marked contrast from models in which citizens accept authoritarian outcomes because on balance these are preferable to the alternative. We illustrate the model’s implications through an analytic narrative that focuses on the current Polish government’s efforts to reform the Polish judiciary.

1. INTRODUCTION

Recent years have witnessed a wave of “authoritarian backsliding” among countries that appeared to be well on their way to being consolidated democracies (Bermeo 2016; Lust and Waldner 2015; Serra 2012). Poland and Hungary provide two vivid examples. Each emerged from behind the Iron Curtain after the dissolution of the Soviet Union, and quickly established reasonably well-functioning democratic regimes. By 1999, they had joined NATO, and five years later became full-fledged members of the European Union. By all appearances, both countries seemed well-ensconced among Western democracies. And yet, only a decade later, each is governed by a party that is arguably chipping away at democratic institutions and norms, raising concerns of creeping authoritarianism that have even been voiced by European Union institutions (Sedelmeier 2014;

Jenne and Mudde 2012; *Revised Statute of the European Commission for Democracy through Law* 2002; of Justice of the European Union 2019a,b; *Sedziowie pod Presja [Judges under Pressure]* 2019; *Lex Super Omnia Association of Prosecutors to the Prosecutor General* 2019). Critics of the Polish government, for example, have argued that its judicial reforms are designed to provide the executive with more direct control over the Supreme Court and the Constitutional Tribunal in order to rule with fewer constraints, to manipulate electoral results, or to persecute political opponents. Similarly, opponents of the Hungarian government allege that its reforms of the constitution and electoral institutions represent an attempt to limit political competition, and to preserve the party's hold on power.

A prominent explanation for such authoritarian backsliding, formalized by Svolik (2018), highlights the significance of polarization and political culture. The logic of this explanation is intuitive: If the ideological alternative offered by the opposition becomes sufficiently unattractive, citizens may be willing to tolerate or even support incumbents with authoritarian tendencies. Polarization, which implies that citizens increasingly view “the other side” as ideologically distant makes it more likely that this condition is met. This condition is also more easily met as citizens are less committed to democratic institutions or processes, i.e., the more heavily ideological considerations weigh in their calculus. Employing experimental evidence from Venezuela, Svolik provides powerful support for this backsliding logic (Svolik 2018).

A key feature of this account is that citizens value ideological proximity more than an institutional commitment to democracy: Accepting a slide into authoritarianism is the price citizens willingly pay in order to achieve ideological congruence, and they pay this price without regret. In this paper, we extend this account to consider whether authoritarian backsliding is also possible when citizens have strong commitments to democratic institutions, i.e., when citizens would not knowingly support a closet autocrat. This question is relevant because there is at least some evidence to suggest that when authoritarian backsliding occurs, it is not always the case that those who support would-be autocrats do so because they lack commitments to democratic values. As an example, consider Poland again. As Figure 1 shows, survey evidence indicates that supporters of the ruling Law and

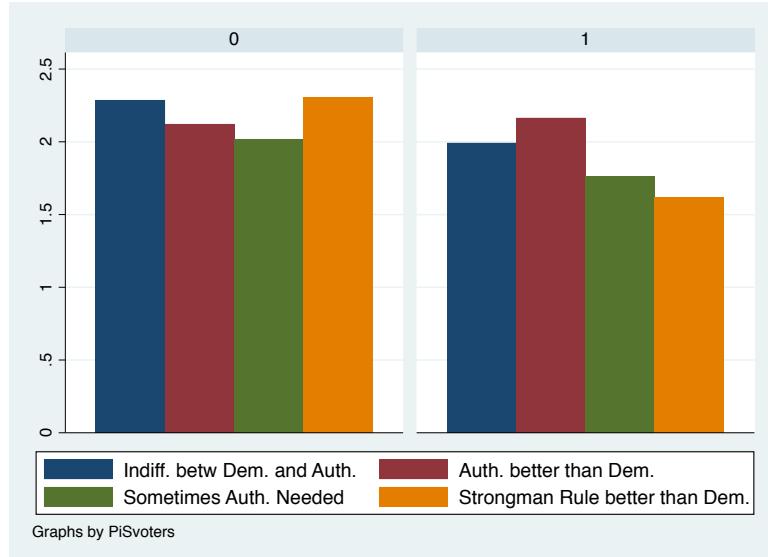


FIGURE 1. Differences in Authoritarian versus Democratic Values among PiS and non-PiS Polish voters

Justice party (PiS) value democracy and the rule of law just as much as those who cast their votes for parties in the opposition.<sup>1</sup> There is almost no difference in responses across the two types of voters, with opposition voters averaging less pro-democratic answers than PiS voters on one of the categories.<sup>2</sup>.

We offer a theory of authoritarian backsliding that reconciles this puzzling combination of popular support for democracy with an ascent to power – through regular electoral channels – of rulers who turn out to be “closet autocrats.” The key to our argument is that authoritarian backsliding is a dynamic process: It occurs over time. Incumbents with authoritarian aspirations usually do not pursue these openly and all at once. Instead, authoritarianism proceeds in incremental steps as current policy choices or institutional reforms are used to lay the groundwork for future authoritarian power grabs (Levitsky and Ziblatt 2018). These dynamics are critical because there typically is disagreement among observers and citizens about the intentions that motivate incumbents to pursue specific policy choices or institutional reforms. Consider the Polish and Hungarian cases

<sup>1</sup> The specific questions asked respondents to what extent they agree or disagree with the following statements “Democracy is superior to any other form of rule,” “For people like me, it does not matter whether the regime is authoritarian or democratic”, “Sometimes Non-democratic rule is better than democratic rule”, “Government by a strong leader is decidedly better than democratic rule.” The answers to the first questions were recoded to match the other three, so that that all four answers increase with pro-democratic values.

<sup>2</sup> To be sure, Svolik does suggest that even in Venezuela voters *when asked* express support for democracy, which is in fact why he goes through the trouble of setting up a survey experiment Svolik (2018)

again. The PiS and Fidesz governments adamantly deny that the goal of institutional reforms is to undermine democratic institutions or norms. Both claim that their reforms are motivated by a desire to enhance democratic accountability, and to remove the residual influence of holdovers from the former communist regime.<sup>3</sup> Voters thus face a potential dilemma: On the one hand, the most direct defense against a “closet autocrat” is to remove the incumbent from power before authoritarianism has advanced too far. On the other hand, voters face genuine uncertainty about whether the incumbent is, in fact, a closet autocrat or is pursuing a sincere policy with no intention of undermining democracy.

In our model, this uncertainty faced by voters regarding the motivations behind institutional reforms adopted by incumbent governments takes center stage. Consistent with other arguments (Svolik 2018; Bermeo 2016), our model suggests that polarization is a critical factor in creating circumstances that allow for “closet autocrats” to retain a hold on power. Additionally, the model demonstrates that the degree of polarization required depends critically on the uncertainty facing voters, and the degree to which voters are concerned about authoritarianism. Finally, and perhaps most importantly, our model shows that the dynamic nature of authoritarian backsliding, and the uncertainty confronting voters, can allow for “closet autocrats” to establish a hold on power even if citizens are fundamentally opposed to authoritarianism – something that is not possible in existing models in which voters knowingly accept an autocrat (Svolik 2018). Consequently, in our model, voters may experience sincere regret when it becomes apparent that they have backed an incumbent who reveals himself to be an autocrat over time. This phenomenon of voter regret is what we will focus on in our empirical implications section.

The paper proceeds as follows. In the next section, we present a simple model that formalizes the theoretical logic. We then derive some empirical conjectures from the model that inform analytical narratives based on the current political situation in Poland and the lead up to the 2015 and 2019 elections. A final section concludes.

---

<sup>3</sup> Indicative of this is Prime Minister Morawiecki’s lecture at New York University earlier this year, where he said “The judges were never appropriately treated in the case of breaching the law before has created the situation. For me it is a situation comparable to the French one in the post-Vichy France. The Prime Minister Michel Debre in 1958 and Charles De Gaulle completely overturned the judiciary system.”

## 2. A THEORY OF AUTHORITARIAN BACKSLIDING

Our model of authoritarian backsliding is designed to capture two key features. First, authoritarian backsliding is typically a dynamic process: It unfolds over time. In particular – as our examples suggest – the initial steps towards authoritarian power grabs are rarely explicitly autocratic. Rather, in most instances, governments intent on pursuing an authoritarian agenda initially engage in institutional reforms that will, subsequently, make it easier to act in a more authoritarian manner. For example, incumbents may reform judicial institutions to loosen judicial oversight, reform media law to reduce journalistic scrutiny, or alter electoral institutions in order to lessen the threat of effective political competition. Once such reforms have taken hold, governments may then be in a position to pursue an autocratic agenda more openly by prosecuting political opponents, manipulating the media, or raising the bar for removing them from power.

This dynamic aspect leads directly to the second key feature: Citizen uncertainty over incumbent motivations. Generally, governments who initiate institutional reforms that can help to pave the way for authoritarian power grabs do not motivate them with autocratic goals; instead, they offer justifications for reforms that emphasize a commitment to democratic values, and claim that reforms are merely intended to facilitate “ordinary” political decisions. Consider Poland again: The Polish government explicitly argued that judicial reform was necessary in order to reduce the hold-over power of communist-era judges, and to strengthen Polish democracy by empowering democratically-elected Polish governments vis-a-vis an authoritarian legacy. Even if they are opposed to authoritarianism, citizens – especially those who are sympathetic to the ideological position of an incumbent – are thus confronted with uncertainty. Is the incumbent committed to democracy, and pursuing reforms in an effort to advance policy goals with which the citizen may agree? Or is the incumbent a “closet autocrat” who will use the room afforded by reforms not only to pursue a policy agenda with which the citizen agrees, but also to undermine democracy in other ways that the citizen opposes?

To capture these aspects, we employ a two period, incomplete information model played among three players: A representative citizen, two types of incumbents, and an opposition. The two period

set-up allows us to model, in the simplest way possible, the dynamic aspect of the process we are considering. Specifically, we assume that in the first period, an incumbent government chooses an institutional reform that will define the actions available to the government in the second period. To model citizen uncertainty, we assume that there are two types of incumbents: One who has autocratic intentions, and one who does not. A representative citizen is uncertain about the type of incumbent in office. Having observed the first-period institutional reform, the citizen updates her belief about the nature of the incumbent's intentions, and chooses whether to reelect the incumbent, or to replace him with the opposition. In the second period, the government-in-office chooses actions in light of its own preferences, and the maneuver room afforded by the institutional reform undertaken in period 1.

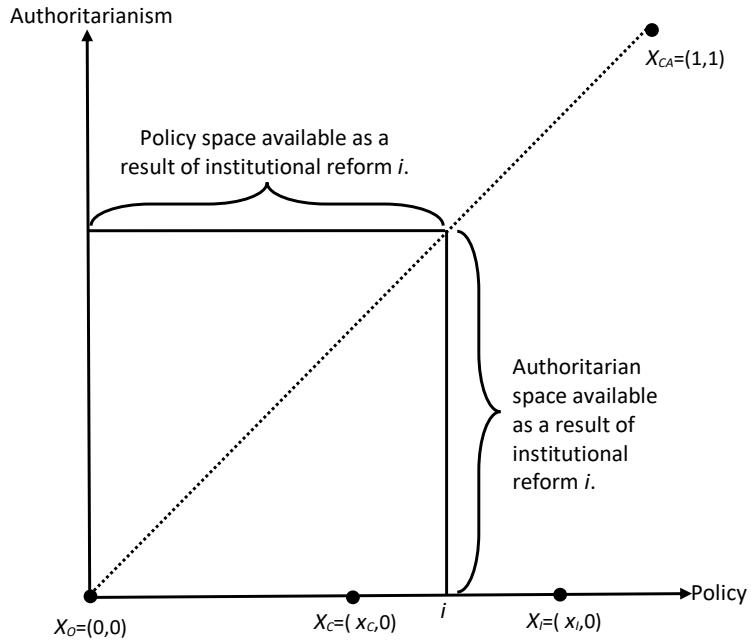


FIGURE 2.

We model the connection between the first-period institutional reform and the actions available to second-period governments in a simple way that captures – at an intuitive level – the fact that institutional reforms can enable both “ordinary” policy goals as well as potential autocratic power grabs. Specifically, the first-period incumbent chooses an institutional reform  $i \in [0, 1]$ . This institutional reform defines a two-dimensional space for the second period,  $[0, i] \times [0, i]$ , as

illustrated in Figure 2. One dimension of this space is an ordinary policy dimension. The other captures authoritarian power grabs. Put differently, institutional reform in the first period enables the second period government to engage in two different decisions: It can make use of the room created by institutional reform to pursue ordinary policy objectives. But it can also take advantage of the opportunities created by the institutional reform to engage in an authoritarian power grab. To return to our example, reforming the Polish judiciary to reduce the influence of communist-era judges may allow a government to engage in more effective “ordinary” policymaking (say, economic reform), but it may also enable it to make authoritarian moves (say, engage in politically-motivated prosecutions of the opposition). We capture this feature by assuming that in the second period, the incumbent takes an action,  $\{p, a\} \in [0, i] \times [0, i]$  that has (potentially) a policy as well as an authoritarian component.

To capture the representative citizen’s uncertainty regarding the motivations of incumbent governments, and the electoral choice confronting the citizen, we assume that there are two types of incumbents. An “ideological” incumbent is purely policy-motivated, and has no interest in authoritarian power grabs. The ideal point of this incumbent is given by  $X_I = \{x_I, 0\}$ , where  $x_I \in (0, 1)$ . In words, the ideologue has a policy ideal point of  $x_I$  but no preference for authoritarianism. The second type of incumbent is a “closet autocrat” whose ideal point is given by  $X_{CA} = \{1, 1\}$ . In words, the closet autocrat prefers to move in the policy dimension, and also wants to undertake authoritarian power grabs. The citizen’s ex ante belief that the incumbent is a closet autocrat is given by  $Pr(\text{Closet Autocrat}) = \alpha \in (0, 1)$ . Finally, the citizen faces an electoral choice between the incumbent (over whose type she is uncertain) and an opposition party whose ideal point is given by  $x_O = \{0, 0\}$ , i.e. the opposition is ideologically to the left of either incumbent type, and it does not want to engage in authoritarian moves.<sup>4</sup>

We assume that the citizen’s ideal point is given by  $X_C = \{x_C, 0\}$ , where  $x_C \in (0, 1)$ . Thus, the citizen is purely policy-motivated (and favors institutional reforms in so far as they enable policy

---

<sup>4</sup> The assumption that the opposition’s ideal point is at  $\{0, 0\}$  is without loss of generality; as we explain below, allowing the opposition to become less moderate, i.e., moving its ideal point below 0 can only exacerbate the threat of authoritarian backsliding.

choices the citizen favors). But she has no authoritarian tendencies, and prefers that the incumbent not engage in authoritarian power grabs. Specifically, we assume that the citizen's preferences over the government's second-period actions are given by

$$(1) \quad U_C(\{p, a\}) = -(p - x_C)^2 - \beta a^2$$

where  $\beta$  is the weight attached by the citizen to authoritarian moves. Because we are interested in authoritarian backsliding when citizens are intrinsically opposed to authoritarianism, we assume that  $\beta \geq 1$ , which ensures that the citizen always prefers the opposition to the closet autocrat if both could implement their ideal points (even if the citizen's ideal point approaches  $\{1, 0\}$ ).<sup>5</sup>

The final element of the model concerns the preferences of officeholders, i.e., the incumbent and the opposition. We assume that these actors are motivated by three concerns. The first is that they place a value  $b \geq 0$  on being in office. Second, they have *instrumental* outcome concerns: They would prefer the second-period outcomes (both in terms of policy and authoritarianism) to be as close as possible to their ideal point. Finally, incumbents have *expressive* concerns: They prefer that the policies they back publicly when in office are closer to their sincere preferences rather than further away. Substantively, one way to think about these expressive preferences is in terms of "branding:" Parties represent a set of policies they are committed to and that form the basis of their appeal. All things being equal, they prefer to announce or back policies that are consistent with these underlying commitments, and dislike those that require them to compromise the party's brand. In the extreme, such expressive preferences may even lead a party to forego being in government, if governing requires it to back policies that are too unpalatable. The Polish PiS and Hungarian FiDesz party can serve as potential examples: For years, both consistently professed positions that kept them in opposition rather than compromising in ways that might

---

<sup>5</sup> This differentiates our approach from Svolik's (2018) model in which citizens vote for the autocrat because the ideological profile of the opposition is not acceptable to them.

make them politically viable but dilute their brand.<sup>6</sup> Thus, suppose the ideological incumbent is reelected and chooses action  $\{p, a\}$  in period 2. His payoff is given by:

$$(2) \quad U_I(\{p, a\}) = -((p - x_I)^2 + a^2) - \theta((p - x_I)^2 + a^2) + b$$

where  $\theta \geq 0$  is the weight attached by the ideologue to expressive preferences. The first term in the expression captures the outcome concerns, the second the expressive concerns, and the final term the office motivations.

Similarly, the payoff of the closet autocrat for choosing action  $\{p, a\}$  is given by:

$$(3) \quad U_{CA}(\{p, a\}) = -((p - 1)^2 + (a - 1)^2) - \theta((p - 1)^2 + (a - 1)^2) + b$$

Summarizing, our model is defined by the following sequence of play:

- (1) Nature chooses the type of incumbent, with  $Pr(\text{Closet Autocrat}) = \alpha \in (0, 1)$ .
- (2) Period 1: The incumbent government chooses an institutional reform  $i_j \in [0, 1]$ , where  $j \in \{I, CA\}$ .
- (3) Period 2: The citizen updates beliefs about the type of the incumbent, and either reelects the incumbent, or replaces him with the opposition.
- (4) Period 2: The second-period government chooses an action  $\{p, a\}_j \in [0, i] \times [0, i]$  (where  $j \in \{I, CA\}$ ), consisting of a choice in the policy dimension and – potentially – an authoritarian move.
- (5) The game ends, and payoffs are collected.

A suitable solution concept is Perfect Bayesian Equilibrium, which requires that players' strategies are sequentially rational, and that they update their beliefs (in our case, about the type of incumbent) in accordance with Bayes' rule along the equilibrium path. We reserve a full statement

---

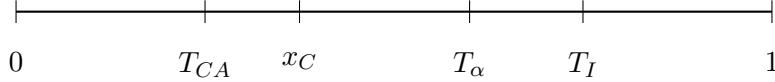
<sup>6</sup> Note that we assume that policymakers have no intrinsic preferences over the first-period institutional reform: Their preferences over institutional reform are induced by their preferences over second-period outcomes. That said, the model's results are robust to incorporating the possibility that policymakers have direct preferences over the institutional reform as well.

of equilibria and proofs to the formal appendix. Here, we focus on presenting the intuition and substantive implications underlying the equilibria.

**2.1. Strategies.** To understand the intuition underlying the model’s equilibria, and its implications for authoritarian backsliding, it is useful to begin by considering the core logic confronting the key players: the two types of incumbents and the representative citizen. We begin with the citizen’s electoral decision. From the citizen’s point of view, the incumbent’s first-period institutional reform is significant for two different reasons. Most obviously, it defines the maneuver room for the second period government. As such, there are potentially two countervailing considerations. On the one hand, the reform can open up room for second-period governments to implement policies the citizen favors. On the other hand, the reform can also enable authoritarian moves, which the citizen opposes. Put differently, the citizen may be sympathetic to reforms to the extent that she is convinced that these reforms will allow for policy choices she agrees with. But she may also be concerned that reforms preserve sufficient constraints on potential autocrats. The second significant aspect of the institutional reform from the citizen’s point of view is that reform can act as a signal of the incumbent’s type; that is, the nature of the reform may provide the citizen with information about whether the incumbent has autocratic goals or not.

The representative citizen’s strategy is characterized by a threshold that makes the electoral decision contingent on the citizen’s (updated) beliefs about the nature of the incumbent, and the extent of the institutional reform that has been adopted. Suppose, for example, that the citizen – having observed the first period institutional reform – has become convinced that the incumbent is an ideologue, and will not exploit the reforms in order to make an authoritarian move. In this case, the citizen’s reelection decision simplifies to a straightforward ideological choice: She reelects the incumbent if the incumbent’s first period reform implies a second period policy that is closer to her ideal point than the opposition, and votes for the opposition otherwise. As is intuitive, this threshold is given by  $T_I = 2x_C$ . We illustrate this threshold in Figure 2, which maps out the citizen’s strategy.

FIGURE 3. Illustration of Citizen Threshold Strategy



In contrast, now suppose that – having observed the first period institutional choice – the citizen believes that the incumbent is a closet autocrat. This knowledge puts the citizen in a quandary: On the one hand, the policy choice of the closet autocrat may be ideologically more palatable than the alternative offered by the opposition. On the other hand, the citizen is opposed to the fact that the closet autocrat will also exploit the maneuver room created by the institutional reform to engage in authoritarian power grabs. The citizen resolves this tension by adopting a strategy under which she will reelect a known closet autocrat *only if* the autocrat chooses a first period institutional reform that is so moderate as to significantly constrain future authoritarian moves. This threshold is given by  $T_{CA} = \frac{2x_C}{\beta+1}$ . Note that – as illustrated in Figure 2 – this threshold is *below* the citizen’s own ideal point: The citizen is trading ideological proximity for constraints on future authoritarianism. Note also that this threshold depends on  $\beta$ , the degree to which the citizen is concerned about the rule of law. As this concern looms larger ( $\beta$  increases),  $T_{CA}$  moves left, implying that the citizen demands more and more significant constraints on future behavior before she is willing to reelect a known closet autocrat.

Finally, suppose that the citizen remains uncertain about the incumbent’s type, i.e., the incumbent’s authoritarian leanings, after the institutional reform. Once again, there is a threshold ( $T_\alpha = \frac{2x_C}{\alpha\beta+1}$  in Figure 2) such that the citizen reelects the incumbent only if the first period reform fall below it.<sup>7</sup> As is intuitive,  $T_\alpha$  falls between  $T_I$  and  $T_{CA}$ , and depends critically on the citizen’s ex ante belief that the incumbent is a closet autocrat ( $\alpha$ ) and the degree to which she is concerned to maintain the rule of law ( $\beta$ ). As she becomes more and more certain that the incumbent is likely to be a closet autocrat ( $\alpha$  goes to 1), this threshold converges on  $T_{CA}$ : The citizen requires more stringent constraints on the incumbent to limit potential authoritarian moves in the second period.

<sup>7</sup> Note that in a more general formulation, the citizen’s updated belief is given by  $\gamma$ , and the threshold is  $T_\gamma = \frac{2x_C}{\gamma\beta+1}$ . Since – as we show below – all equilibria in which the citizen remains uncertain over the incumbent’s type are fully pooling,  $\gamma = \alpha$ , which is why we write the threshold in terms of the ex ante belief.

In contrast, as she becomes more convinced that the incumbent is an ideologue ( $\alpha$  goes to 0), this threshold converges on  $T_I$ . (Note that this implies that  $T_\alpha$  may be to the right or left of  $x_C$ .)

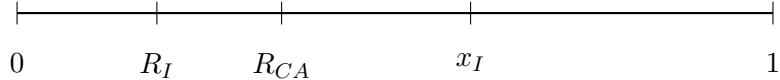
Having worked through the logic of the citizen's reelection strategy, consider the two types of incumbents as they decide on institutional reform in the first period. The key issue confronting the incumbent is that the institutional reform he adopts has two separate effects: It determines the citizen's reelection decision, and it defines the maneuver room the incumbent has in the second period (if reelected). There are clear incentives to be re-elected, since office provides intrinsic benefits as well as the opportunity to make policy (including the possibility of pursuing an autocratic agenda). This implies that there is value in proposing an institutional reform that will result in reelection, even if this reform is not (from the incumbent's point-of-view) ideal. At the same time, for the incumbent, there are expressive costs associated with governing in ways that are not in line with the incumbent's preferences, that is, to have to adopt a policy package in the second period that diverges from the incumbent's ideal. As a result, there is a limit to the incumbent's willingness to adopt an institutional reform that will result in his reelection. If the institutional reform that ensures reelection is so modest as to constrain the second period policy choice too much, the incumbent prefers to leave office and tolerate the opposition's ascent to power.

These considerations imply that the two types of incumbents adopt a threshold strategy. For each incumbent, this threshold marks the most modest institutional reform the incumbent is willing to adopt in order to be reelected. Reforms to the left of these thresholds are so constraining as to make returning to power sufficiently unattractive. These thresholds are summarized in Table 1 (along with the thresholds imposed by the citizen), and Figure 3 provides a graphical illustration, where  $R_{CA}$  denotes the constraint imposed by the closet autocrat and  $R_I$  denotes the threshold imposed by the ideologue. Note that  $R_{CA} > R_I$ . This is a critical feature that plays a key role in the equilibria. The intuition behind this ordering should be clear: For the closet autocrat, institutional constraints on power are more costly than for the ideologue, since the closet autocrat is constrained in two rather than just one dimension. As a result, the closet autocrat is quicker to bristle at an institutional constraint.

TABLE 1. Summary of thresholds

Threshold	Imposed by	Explanation
$T_{CA} = \frac{2x_C}{\beta+1}$	Citizen	Citizen will reelect known closet autocrat who sets reforms below this threshold.
$T_I = 2x_C$	Citizen	Citizen will reelect known ideologue who sets reforms below this threshold.
$T_\alpha = \frac{2x_C}{\alpha\beta+1}$	Citizen	If the citizen is uncertain about the incumbent's type, she will reelect an incumbent who sets reforms below this threshold.
$R_{CA} = 1 - \frac{\sqrt{2+b}}{\sqrt{2(1+\theta)}}$	Closest autocrat	Closest autocrat will only set reforms at or above this threshold to get reelected.
$R_I = x_I - \frac{\sqrt{b+x_I^2}}{\sqrt{1+\theta}}$	Ideologue	Ideologue will only set reforms at or above this threshold to get reelected.

FIGURE 4. Illustration of Incumbent Threshold Strategy



**2.2. Equilibria and Substantive Implications.** The model's equilibria, which are unique for any combination of parameters, depend on the relative position of the thresholds that characterize the citizen's and the incumbents' strategies. The equilibria are described in the formal appendix. Here, we focus on the substantive implications, which are most readily explored by grouping the seven equilibria into four types. Three of these are separating equilibria in which the ideologue and the closet autocrat adopt different institutional reforms in the first period, allowing the citizen to learn the type of incumbent she is facing. What distinguishes these equilibria is the manner in which the citizen responds to the institutional reforms that are adopted:

- (1) Type 1: Given the institutional reforms adopted, the citizen replaces both types of incumbent with the opposition.

- (2) Type 2: Given the institutional reforms adopted, the citizen only reelects the ideologue and replaces the closet autocrat with the opposition.
- (3) Type 3: Given the institutional reforms adopted, the citizen chooses to reelect both the closet autocrat and the ideological incumbent.

In addition to these separating equilibria, there also exist two pooling equilibria in which the ideologue and the closet autocrat adopt the same institutional reform in the first period, and the citizen remains uncertain about the type of incumbent. In these pooling equilibria, the citizen reelects the incumbent (Type 4).

To discuss the substantive interpretation of these equilibria, and the implications for authoritarian backsliding, it is useful to graph the location of these equilibria in the parameter space. We do so in Figure 5, which shows where each of the types of equilibria occurs as a function of the ideal point of the ideological incumbent (plotted on the  $x$ -axis) and the ideal point of the citizen (plotted on the  $y$ -axis). The solid lines indicated the partition of the four types of equilibria. The light dashed lines indicate the partition of the equilibria that fall under each type. In addition, the figure displays, for each equilibrium, the institutional proposal made by the ideological incumbent ( $i_I$ ) and the closet autocrat ( $i_{CA}$ ) in the first period. (For example, the northwestern corner of the space features the Type 3 separating equilibrium, in which the ideological incumbent proposes  $i_I = x_i$ , the closet autocrat proposes  $i_{CA} = T_{CA}$ , and both types are reelected as a result.)

We discuss what we can learn from the model about the phenomenon of authoritarian backsliding through a series of implications, using Figure 5 as a guide. We begin with the equilibrium that most closely resembles the logic established by Svolik (2018), and then consider the consequences of extending this logic to a dynamic setting in which citizens face uncertainty over the intentions of incumbents.

**Implication 1.** *In an environment in which polarization in society far surpasses polarization of elites in the sense that the citizen has considerably more extreme preferences than the ideological*

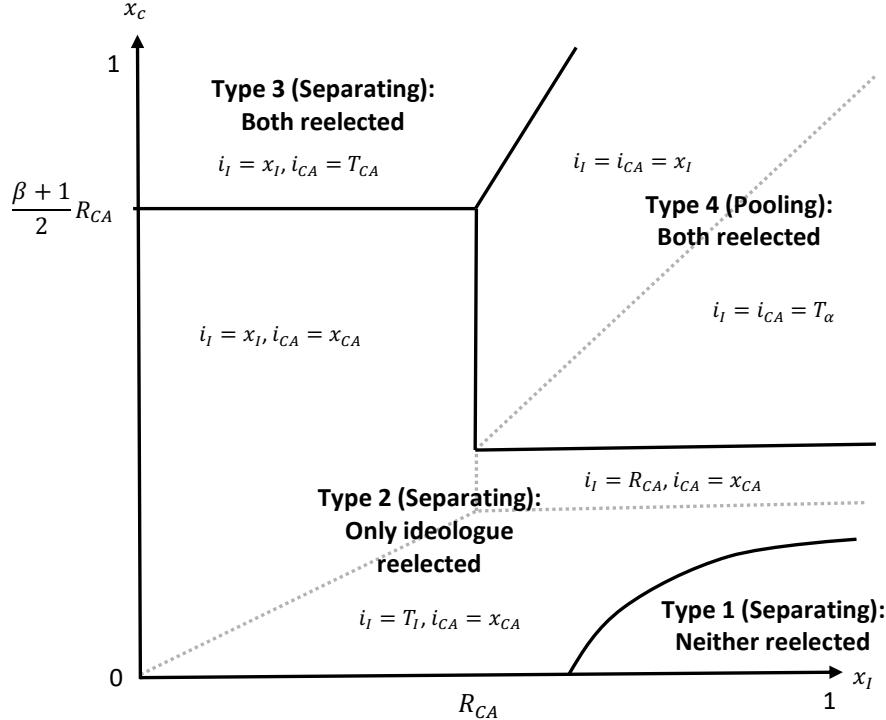


FIGURE 5.

incumbent and the opposition, the two types of incumbents separate and both are reelected. Authoritarian backsliding occurs if the incumbent is the closet autocrat, and the citizen knowingly opts for this alternative.

This implication captures the separating equilibrium in the northwestern part of the figure, and describes the situation modeled by Svolik (2018). The citizen's preferences are so extreme, relative to the "mainstream" parties, that she prefers to reelect the incumbent *even when she knows* that the incumbent is a closet autocrat. Put differently, in this equilibrium (as is the case in the Svolik account), there is no room for citizen regret; the citizen opts for reelecting a closet autocrat knowing full well what the implications of doing so are. As is intuitive, the conditions that allow for this equilibrium become more restrictive as citizens are more concerned to prevent authoritarian backsliding (i.e., as  $\beta$  increases). Against this baseline, we are now in a position to consider the consequences of our extension of the Svolik account.

**Implication 2.** *In a polarized political environment, in which the preferences of the ideological incumbent and the citizen are sufficiently far to the right of the opposition, the two types of incumbents adopt the same reform in period 1, and are reelected by the citizen. Authoritarian backsliding in period 2 is possible. When it occurs, the citizen – ex post – regrets her electoral choice.*

This implication captures the pooling equilibria in the northeastern part of the figure. As the political environment becomes more polarized, i.e., as the citizen and the ideological incumbent move to the right, two dynamics begin to play out. The first is the logic identified in Svolik's (2018) model: Replacing the incumbent with the opposition becomes less palatable for the citizen. Even if there is some possibility that the incumbent is an autocrat, the citizen is more willing to reelect the incumbent in order to obtain policies that are ideologically proximate. The second dynamic concerns the distinct element of our model: Citizen uncertainty about the type of incumbent. As the citizen is moving to the right, the institutional reform that the ideological incumbent can adopt and still be reelected also shifts to the right. The consequence is that the cost to the closet autocrat of playing the “wolf in sheep’s clothing” declines. Once the citizen and the ideologue are sufficiently far to the right, the closet autocrat is willing to adopt the same institutional reform as the ideologue – in effect “masquerading” as an ideologue. Once this happens, the citizen – unlike in the separating equilibria that surround this region of the parameter space – is no longer sure about the nature of the incumbent. But given that her preferences are relatively far to the right, she is willing to take a gamble, and chooses to reelect. If the incumbent turns out to be the closet autocrat, he will exploit the room created by the institutional reform to pursue authoritarian policies in period 2. Authoritarian backsliding occurs.

Critically, when such backsliding does occur (i.e., when the incumbent reveals himself to be a closet autocrat upon reelection), the citizen feels genuine regret in the sense that she would – given the ideological and autocratic policy choices being made by the closet autocrat in the second period – prefer to “swap out” the closet autocrat for the opposition. But it is now too late. Such an outcome is not possible in the Svolik account (or the equilibrium discussed in Proposition 1), in which citizens knowingly accept authoritarianism in return for ideological proximity.

**Implication 3.** *As the citizen is (a) less concerned about authoritarian backsliding ( $\beta$  decreases), or (b) believes that it is less likely that she is facing a closet autocrat ( $\alpha$  approaches 0), the area of the parameter space covered by the pooling equilibrium increases. As a result, the level of polarization required for the pooling equilibria decreases. Authoritarian backsliding is easier to achieve.*

This implication highlights that there is an interactive relationship between the level of polarization (in terms of the location of  $x_C$  and  $x_I$ ) necessary to open the door to authoritarian backsliding, and the degree to which the citizen is concerned about the threat of authoritarianism. As the citizen believes that it is less likely that she is confronting a closet autocrat ( $\alpha$  is low), or as she cares less deeply about preventing autocratic moves ( $\beta$  approaches 1), the citizen is more willing to take a gamble on reelecting an incumbent about whose intentions she is uncertain ( $T_\alpha$  moves to the right). And as the citizen becomes more willing to reelect incumbents even if they adopt more extreme reforms, closet autocrats perceive an opportunity to masquerade as ideologues by adopting institutional reforms that will allow them to be reelected. In other words, political environments in which citizens believe the threat of authoritarianism is low, or in which citizens have shallow commitments to democratic institutions, are fertile ground for closet autocrats to begin infiltrating a democratic polity.

So far, we have discussed two routes, suggested by the model, by which authoritarian backsliding may occur. But the model also has implications for the conditions that prevent authoritarian backsliding. We summarize these in our last implication:

**Implication 4.** *In a “Downsian” environment, in which the ideological incumbent, the opposition, and the citizen are all located sufficiently close to each other, the two types of incumbents separate by proposing different institutional reforms, and only the ideologue is reelected. No authoritarian backsliding occurs.*

This implication is illustrated by the southwestern part of the figure – an environment we refer to as “Downsian” in the sense that the opposition (with an ideal point at 0) and the ideological incumbent (with an ideal point of  $x_I$ ) are relatively close together, and both are located close to

the citizen's ideal point. In this environment, when the ideologue is closer to the citizen than the opposition is, the ideologue can pursue his preferred reform in period 1 and be reelected. If the ideologue is further from the citizen than the opposition is, he moderates his proposed reform, constraining himself in order to be reelected. In either case, the reform proposed by the ideological incumbent is so moderate that the closet autocrat is not willing to emulate the ideologue in order to be reelected. Interpreted more broadly, these equilibria suggest that political environments in which the “mainstream” parties (the ideologue and the opposition) are sufficiently centrist (in the sense of being close to the pivotal voter) isolate closet autocrats, and make them politically unviable. Autocrats cannot “infiltrate” the political system because the appeals that they are willing to run on will – given the nature of the political alternatives available – expose them, and lead to rejection by voters. Consequently, no authoritarian backsliding occurs.

**2.3. Summary.** Before specifying the empirical implications of our model that are the focus of the analytic narrative of the next section, it is useful to summarize the central logic of the argument we have laid out. The aim of the model is to contribute to understanding how authoritarian backsliding can occur in contexts in which citizens (at least pivotal citizens) are, at bottom, *opposed* to creeping autocracy. The model is designed to capture two factors that we regard as significant. The first is that there is a dynamic element to authoritarian backsliding: Governments take actions at one time (often in the form of institutional reforms) that may enable authoritarian moves in the future. The second is that there are typically competing interpretations of these actions: Citizens face uncertainty about whether the incumbent government is a closet autocrat intent on laying the groundwork for future authoritarian moves, or an ideologue who may have ideologically more extreme preferences but is otherwise committed to the rule of law.

The central insight of the model comports well with intuition: Increasing polarization – both of the electorate, and of political elites – is a central driver of the potential for authoritarian backsliding. There are two reasons for this. The first – an insight originally developed by Svolik (2018) – is that polarization makes electoral punishment of the incumbent more difficult by making the opposition less attractive to potentially pivotal citizens. The second feature is closely related

to this, and derives from the uncertainty that voters face about the incumbent's intentions (an element not addressed in Svolik's model): By making more extreme policy choices electorally viable, polarization opens the door to closet autocrats who may enter the political system because they see an opportunity to begin laying the groundwork for authoritarian power grabs while retaining power. Being uncertain about the incumbent's ultimate intentions, citizens – given sufficient polarization – may be willing to reelect an incumbent only to discover afterwards (and with regret) that the incumbent is in fact a closet autocrat.

The second key insight of the model is that how much polarization is required in order to give rise to these dynamics depends critically on voter beliefs about the nature of the incumbent they are facing, and the extent to which voters are concerned with preventing authoritarianism. The less heavily concerns over autocracy weigh on citizens, or the more citizens are convinced that the incumbent is an ideologue, the more likely it is that even moderate levels of polarization can set off the process of authoritarian backsliding. This second insight has clear implications for established democracies. In such settings, citizens have little experience with autocrats and may be less attuned to the possibility that incumbent actions are motivated by authoritarian intentions. This may be particularly true when closet autocrats do not present themselves in the guise of newly created parties, but instead emerge as a faction within an existing party with long-established democratic credentials. In such scenarios, citizens may – mistakenly – be too confident that they are dealing with an ideological incumbent. In short, it is exactly in established democracies, where citizens believe that incumbents are ideological rather than closet autocrats that such closet autocrats would have an easier time coming to power.

### 3. ANALYTICAL NARRATIVE

The model we have presented is highly simplified, and focuses on a small number of actors and decisions in order to put one mechanism that may contribute to authoritarian backsliding into starker relief. This calls for a multifaceted approach to connecting the logic of the model to “real world” situations. We adopt such a multifaceted approach with an analytic narrative built around the last few years in Poland, since the 2015 elections.

Analytic narratives have been successfully used to illustrate predictions from formal models in the past by applied theorists (Levi 2004; Bates 2007; Vanberg 2000; Nalepa 2010; Lorentzen 2014; Lorentzen, Fravel and Paine 2015; ?). The equilibrium of a formal model is not just a static state of the world, but an entire plan of action for each player of the game, even for contingencies that may never arise. Testing such a set of rationally consistent responses even to counterfactual states of the world will require uncovering actors' subjective beliefs. A qualitative case study can help unveil the reason certain actors chose the actions they did just as a semi-structured elite interview can reveal that these actors are doing x in order to avoid y, offering the researcher insight what his behavior would have been in states of the world that will never be observed.<sup>8</sup> Historical memoirs and documents in archival repositories can similarly disclose the reasoning behind choosing certain actions rather than others (Goemans and Fey 2009).

While in principle, there is no obstacle to collecting such data on a larger scale, such research places proportionately greater resource demands on scholars. Hence the trade-off associated with qualitative case studies is that the number of independent observations is necessarily smaller. This does not mean, however, that the selection of qualitative cases is not guided by the research design.

Our analytic narrative centers on the pooling equilibrium in which the voters reelect the incumbent because they believe the odds are good enough to secure an ideological incumbents who will implement policy proximate to their ideal point. The Polish case is an ideal interpretation of this equilibrium, because the incumbent has just completed a series of institutional reforms, most of which concern the judiciary. These judicial reforms can be interpreted as paving the way for an authoritarian power grab, but they can also be interpreted as attempts to make the judiciary more effective or to hold accountable for political actions judges and prosecutors with careers that started in the pre-1989, communist era.

Limiting data collection only to variables that operationalize the equilibrium outcome could not offer a complete test of the pooling equilibrium of our model (and this argument is somewhat more technical). Whereas in the case of separating equilibria, where each type of the “informed” player

---

<sup>8</sup> In game theoretic terms, this is insight into off-the-equilibrium-path behavior (Gailmard and Patty 2012).

selects a different action, a quantitative test can be applied to uncover a correlation between type and action chosen, this cannot be done with pooling equilibria, where both types of the informed player select *the same* action. Demonstrating the occurrence of a pooling equilibrium, hence, in contrast to a separating equilibrium requires a qualitative approach, because one can only show that the subjective beliefs of the uninformed party are consistent with pooling: namely, that they cannot discriminate between the types, after observing the action of the informed player.

The dilemma facing Polish voters in the fall of 2019, as the country nears national elections to the Sejm and Senate (the lower and upper houses of the Polish legislative assembly) reflects, as we show, the uncertainty of the pooling equilibrium. These elections will offer Polish voters the first opportunity to reelect the incumbent even though his actions are consistent with the behavior of an ideologue *and* a closet autocrat.

Moreover, to show how critical subjective beliefs here are, we supplement the qualitative case study with a large-n test that tries to manipulate individual beliefs of PiS voters to verify if this produces a corresponding change in choice. We propose an experiment nested within a survey to show that as the beliefs of citizens regarding the type of incumbent they are facing change in favor of closet autocrat they will become less likely to support the PiS incumbent. The experiment is inspired by a similar experiment conducted in the proximity of the Hungarian 2014 elections, which for that country, was the opportunity to reelect Victor Orban, the incumbent leader of Fidesz, as Prime Minister (Ahlnquist et al. 2018).

The research design of this experiment will follow a the presentation of the Polish institutional reforms as a compelling illustration of the ambiguity of the first period action that could come either from an ideological incumbent or a closet autocrat.

**3.1. Policy choices in first period: institutional reforms.** The Polish Law and Justice party (PiS) emerged victorious in the October 2015 parliamentary elections, after sitting in the opposition benches for 8 years. Even though it won only a minority of the vote, it took an absolute majority of seats in the legislature. The victory allowed PiS to establish Poland's first single-party majority cabinet since 1989. The institutional reforms it has implemented in the last four years

concentrate on the judiciary and specifically Supreme Court, National Council of the Judiciary, and the Constitutional Tribunal, in addition to ordinary courts.<sup>9</sup>

The most drastic set of reforms concerned the Supreme Court whose composition was significantly altered by July 2017 legislation. It mandated that judges serving on the Supreme Court who had reached the age of 65 by July 3, 2018, would retire on 4 July 2018, unless they submitted a declaration indicating the desire to continue carrying out their duties and a certificate of good health. The final work authorization, though, was up to Andrzej Duda, the PiS President.<sup>10</sup> The functioning of the court was modified as well with a newly created “disciplinary department,” where 12 newly appointed judges were given posts.<sup>11</sup> The prerequisites for holding a Supreme Court seat were lowered to a minimum of 12 years of experience in a regional court. In Poland, the Minister of Justice simultaneously holds the position of Prosecutor General. Consequently, as a result of the reforms the ruling majority is able in every single court case, to appoint both the prosecutor *and* presiding judge.

Potentially, these measures violated at least two articles of the 1997 Constitution (181 and 182), but neither of them received scrutiny from the Constitutional Tribunal.

Simultaneously, the PiS-led parliamentary majority worked on legislation changing composition of the National Council of the Judiciary (KRS), an independent body which makes recommendations of who should be appointed as a judge (including to the Supreme Court) and also initiates disciplinary action against members of the judiciary.<sup>12</sup> Prior to the bill put forward by PiS, the KRS was made up exclusively of judges. The new PiS legislation fired the existing members and replaced them with political appointees.

---

<sup>9</sup> Poland’s system of courts separates the function of the Supreme Court from the constitutional court, called the Constitutional Tribunal. In addition to being the court of appeal for lower-level courts, the Supreme Court is responsible for determining the validity of elections and nationwide referenda. Meanwhile, the Constitutional Tribunal serves as a lateral institution to the Supreme court. The Tribunal is the equivalent of the Supreme Court in the US and constitutional courts in other countries, except that it only deals with the constitutionality of legislation passed in the Sejm and the constitutionality of legal norms applied in decisions of lower level courts.

<sup>10</sup> In the initial version of the law, the Minister of Justice was handed discretion over which judges could remain in office, but President Duda vetoed the bill, increasing hopes of rule of law -oriented citizens that he had assumed the role of “enforcer of the constitution.” Yet, his veto proved to be more of inside-PiS power struggle with the Justice Minister, who was Duda’s rival within PiS

<sup>11</sup> Technically, this brought the total number of judges in the Court to 43, but effectively, 31 would be doing the work that used to be performed by 87.

<sup>12</sup> The final disciplinary decisions are then carried out by a special Ombudsman for Discipline

The comprehensive reforms did not omit lower courts either. Judges over 65 were forced to retire, unless they received an exemption from the Minister of Justice. Permission would only be granted following vetting by a special commission. The reforms also allowed the government to initiate disciplinary action against judges because the parliament already had been granted the authority to appoint members of the National Council of the Judiciary, the body responsible disciplining judges. A document entitled “judges Under Pressure” *Justitia*, details the disciplinary proceedings initiated by the disciplinary chamber of the reformed Supreme Court. All judges facing disciplinary action had in one way or another objected to the changes implemented following PiS’s reform. Many of the judges facing disciplinary action are themselves members of *Justitia*, an association of judges dedicated to values of rule of law and court independence. (*Sedziowie pod Presją [Judges under Pressure]* 2019). Not only judges, but also prosecutors have been facing harassment, although the actions of prosecutors cannot be subject to disciplinary action by the Supreme Court, the Prosecutor General—a post also controlled by PiS—can take measures to make their personal lives very hard, for instance by moving them from one city to another, even across Poland. Such repercussions faced a couple of prosecutors whose post were relocated in this way (Mariusz Krason and Monika Frackowiak).<sup>13</sup>

Just as in the case of reforms lowering the retirement age of Supreme Court judges, this action provoked no response from the Constitutional Tribunal. Why not? Possibly, because the constitutional court was already controlled by PiS as a result of reforms to the Tribunal implemented within a few months of the formation of the PiS government. The reforms were triggered by crisis that emerged when President Andrzej Duda had refused to swear in three justices to the court who supposed to replace judges whose terms were running out. The three however, had been elected by a parliament still controlled by the Civic Platform (PO), PiS’s predecessor. The Constitutional Tribunal upheld the constitutionality of the vote, but the PiS minister responsible for publicizing the verdict withheld its printing. The three justices joined the bench nonetheless.

---

<sup>13</sup> See for details open letter issued by Lex Super Omnia Association of Prosecutors to the Prosecutor General (*Lex Super Omnia Association of Prosecutors to the Prosecutor General* 2019)

Tensions escalated when PiS using its parliamentary majority elected its own three candidates, who also joined the bench. In response, the Tribunal's Chief Justice, Andrzej Rzeplinski, refused to appoint them to sit on panels. PiS then decided to get rid of Rzeplinski by shortening the terms of all Constitutional Tribunal judges who were older than 65. A PiS loyalist, Julia Przylebska, was appointed to replace him.<sup>14</sup>

The result of these complex reforms was a judicial system controlled from the top down by PiS's ruling majority.

**3.2. Ideologue or wolf in sheep's clothing?** While radical, the institutional reforms implemented by PiS are as consistent with actions of an ideological incumbent as they are with the actions of a closet autocrat. Why would an ideologue embark on a series of judicial reforms?

Here is how President Duda defended the reforms during the anniversary celebrations of the Gdansk Agreements, the 1980 negotiations that took place between Poland's communist government and the independent Solidarity trade union.<sup>15</sup> As crowds of protesters gathered and chanted "Constitution, constitution!", alluding to the fact that the President was falling on the job of enforcing the Basic Law, Duda remarked "You are allowed to protest because I respect the constitution, but rest assured that a majority of people living on the Coast feel threatened by the fact that there are justices on the benches of courts who pronounced sentences on members of the opposition during Martial Law." What Duda was invoking as a justification for the reforms was the incomplete transitional justice project—that is, the unfinished process of settling accounts with members and collaborators of previous communist regime—that always left the judiciary exempt from various lustration and decommunization provisions (Nalepa 2010).

The courts have also been notoriously inefficient with queues in commercial cases to get on the docket stretching into years. This has definitely contributed to lost foreign investment and slower growth than what Poland could have experienced. Thus, increasing accountability of judges could

---

<sup>14</sup> This was a controversial choice for at least two reasons. First, Poland's National Council of the Judiciary had evaluated—and rejected—Przylebska as lacking qualifications for a post in the Appellate Court. What is more, she began her judicial career in communist Poland, making her exactly the kind of judge PiS promised to get rid of.

<sup>15</sup> To many those events were the beginning of the end of communist rule in Europe.

be explained by policy preferences of an ideological incumbent. The ‘P’ in PiS does stand for “law” after all.

At the same time, tinkering with the National Council of the Judiciary and the composition of the Supreme Court carries enormous payoffs to an incumbent who is in fact a closet autocrat. Poland has a public system of financing electoral campaigns: upon clearing the threshold of 3% of the national vote, parties seek reimbursements for campaign expenditures up to 4.04 PLN per vote from the state treasury. Yet to be reimbursed, the applicant’s books must be deemed “in order” by Supreme Court judges. A PiS-controlled Supreme Court could be used to gradually bankrupt opposition parties and eliminate electoral competition to PiS in the future.

Even more directly, the Supreme Court could target directly the most influential opposition party to date—Civic Platform (PO) if Donald Tusk, its former leader, and current European Council president, were put on trial before the State Tribunal.<sup>16</sup> The chief justice of the Supreme Court serves, *ex officio*, as the justice presiding over the State Tribunal. Tusk’s alleged crime is the murdering of Jaroslaw Kaczynski’s twin brother Lech, who was Poland’s president at the time he perished in a plane crash over Smolensk, Russia. According to Jaroslaw Kaczynski, Tusk sabotaged the investigation into the catastrophe and allowed for declaring it an accident much sooner than it was warranted to do so. Finally, gaining steering control over the Supreme Court has allowed Kaczynski to pardon a close ally and associate, Mariusz Kamiński.<sup>17</sup> Mariusz Kamiński is currently chief of the Polish secret service and has just been nominated also as the Minister of Internal Affairs.

In established democracies, legislation that skates the border of compliance with the rule of law, is subject to constitutional review by courts such as the Constitutional Tribunal. However, as we explained above, that court had come under the control of PiS. It is not clear if this reform was carried out deliberately, as had the outgoing government not voted the replacement of as many as 5 of the Constitutional Tribunal justices whose terms were near expiration<sup>18</sup>, perhaps the

---

<sup>16</sup> The State Tribunal is a special judiciary body for assessing the constitutional liability of persons holding the highest state rank. This process could result in criminal punishment and a loss of civil rights.

<sup>17</sup> In a 2015 case, a regional court sentenced of Kaczynski and former head of the Central Anticorruption Bureau, to three years in prison for abuses of power. In November of 2015, within days of assuming office, President Andrzej Duda pardoned him. But the Supreme Court annulled the pardon in March 2017. Barring a reversal, Kamiński would go to prison.

<sup>18</sup> The election of two of the five *was* actually ruled unconstitutional

new PiS government would not have felt provoked to retaliate with sending the Chief Justice of the Court into early retirement when he refused to appoint to cases the PiS elected replacements. Regardless, however, of who “started it first,” it is clear that the body that would under normal circumstances adjudicate between institutional reforms coming from an ideological incumbent and a closet autocrat was unable to come across as impartial.

Similarly ambiguous is the interpretation of the reforms to the lower (so-called “ordinary”) courts. Lower level courts are surprisingly influential due to a peculiarity of the Polish constitution, which allows them to engage in interpreting the constitution when the Constitutional Tribunal is incapable of doing so.<sup>19</sup> This aspect could jeopardize plans of a potential closet autocrat, because even after replacing key constitutional justices of the Constitutional Tribunal, a ruling party could have unconstitutional legislation struck down by ordinary judges. Of course, no ruling party can replace all lower court judges but it can influence their careers with a Supreme Court—the court of appeal for lower-level decisions—that has been brought under the control of the autocrat. With such a politically controlled Supreme Court, any judge interpreting the constitution at odds with PiS risks having the decision reversed. Since frequent reversals undermine judicial careers, few lower-court judges will likely choose this path. And so, to ensure the loyalty of judges in lower level courts, the PiS-led government could following the elections, intensify disciplinary actions against judges deemed in violation of “professional ethics.” But whether it will pursue this path remains unknown.

PiS has also been very careful to avoid the appearance of abusing its power. In the last few months before the elections, PiS Speaker of the Sejm, Marek Kuchcinski, was accused of flying his family members on a government jet around the country, clearly exploiting his privilege as public official. Within days he was forced to resign. At the same time however, PiS may have wanted to replace Kuchcinski with a less independent minded emissary anyway to ensure its executive dominance in the legislature and preserve negative agenda setting power there (Nalepa 2016).

---

<sup>19</sup> Article 178 of the Basic Law

Thus even the dismissal of PiS officials behaving like authoritarian power-grabbers leaves voters trying to discriminate between the rival theories of who their incumbent really is uncertain.

**3.3. The muted response from the EU.** In contrast to Turkey and Venezuela, who have no international organization to sanction clear infringements of rule of law, Poland is a member of the European Union. And so, as early as 2016, Brussels initiated its rule of law procedure into Poland's judiciary reforms, the first time such an investigation had been launched against a member state.<sup>20</sup>

Two elements of the Supreme Court bill sparked the interest of the European Commission: First, the fact the new retirement age applied to judges whose term was in progress; and, second, that judges past the new retirement age would be allowed to serve at the discretion of the executive. Consequently, the Commission brought an “action for failure to fulfill obligations before the Court of Justice.” The case was sent to the Court of Justice of the European Union (CJEU) on October 2018 and further, by order of 15 November 2018, the President of the Court granted the Commission’s request to decide this action under an expedited procedure. This request was granted even *after* PiS had passed amendments to the law in November, 2018.<sup>21</sup>

But EU institutions are slow-moving, particularly as they engage the PiS government allowing it to respond to accusations in writing. In addition, every specific set of sanctions has to be approved by the EU Council, which is made up of cabinet Ministers of all member states. For obvious reason, cabinet ministers are reluctant to do so towards fellow cabinet ministers.

Three hearings of Polish government have taken place as part of this process. During the last one of these, Guy Verhofstadt cited the Venice commission as saying that the judiciary reforms bare “a striking resemblance to methods employed by the Soviet Union and its satellites.”<sup>22</sup> While

---

<sup>20</sup> The EU Treaty’s article 7 allows for the possibility of calling to task a Member State who allegedly breaches the core values of the EU: democracy, human rights, or the rule of law. Most researchers of EU institutions when discussing this article would add: “this would never be used. This is there just for symbolic reasons.” (Joseph H.H. Weiler, moderating NYU panel). Hence it is not a surprise it was never invoked until now.

<sup>21</sup> The gist of the amendment was to reinstate the judges sent to early retirement, yet CJEU continued to be have issues with the law, other parts of which skated on the broader of breaking EU law; in particular, the disciplinary chamber allowing for disciplinary investigations against judges applying the “preliminary reference” clause, which alerts the European Union Court of Justice to legislation potentially in violation of it charter

<sup>22</sup> The Venice Commission (European Commission for Democracy through Law) is a legal advisory body whose primary task is to provide states with legal advice in the form of “legal opinions” on draft legislation or legislation already in force which is submitted to it for examination. It also produces studies and reports on topical issues. Groups of members assisted by the secretariat prepare the draft opinions and studies, which are then discussed

the process of sending other pieces of PiS legislation to the Court of Justice of the European Union (specifically, the establishment of the political KRS and the creation of the disciplinary chamber within the Supreme Court) came to a halt with the approaching of EU elections in May 2019, the European Court of Justice proceeded towards issuing its ruling. As an interim step in this process, on June 24, 2019 the European Court of Justice issued an opinion (an opinion precedes a verdict and typically falls on the same side as the verdict) that lowering of the retirement age of judges in the 2018 bill breached EU law of Justice of the European Union (2019a). Zbigniew Ziobro, the PiS Minister of Justice called the decision incoherent and defended the bill as “the only cure for the cancer which kills Polish judiciary.” Leszek Mazur, meantime, the judge who serving as the president of the National Council of the Judiciary said that were the EUCJ also to rule in conformity with this opinion , it would be infringement of Polish sovereignty.

A final verdict from the Court of Justice of the European Union is expected in the Fall of 2019, most likely right before the parliamentary elections that will give the PiS incumbent a chance to be reelected. It is this window between the EU verdict and the parliamentary elections that we believe is ideal for fielding our survey experiment described in the next section.

The next subsection will corroborate the finding that a change in voters’ beliefs that they are are dealing with a closet autocrat,  $\alpha$ , will decrease their readiness to reelect the incumbent.

#### 4. EVIDENCE FROM SURVEY EXPERIMENT (RESEARCH DESIGN ONLY)

In the section above we explained why PiS and its supporters may believe that all of the institutional reforms they have authored make the courts more efficient, curb corruption of the judiciary and are completing the unfinished decommunization project. At the same time, other voters may be fearful that these reforms are paving the way for a true authoritarian power grab. A situation where the typical PiS voter after observing this incumbent’s actions, is left uncertain about the type of incumbent she is facing is consistent with the pooling equilibrium of our model. Hence, our prediction for the upcoming Polish election would be that PiS will be reelected to govern. And of

---

and adopted at the Committee’s plenary sessions. *Revised Statute of the European Commission for Democracy through Law* (2002)

course, it remains to be seen if following this reelection—corresponding to the second period of our model, where policy choice on the authoritarian dimension is possible—Kaczynski will indeed turn out to be the closet autocrat many believe him to be.

But this research design alone would only allow us to count Poland as a single case that may or may not be consistent with our predictions. We can do better if we allowed every one of a sample of representative PiS voters to make a choice in response to information that should change their beliefs. The kind of information that may have this effect, would have to be authored by a body independent of the government *and* the opposition. Arguably the Court of Justice of the European Union is such a body. As the narrative above explains, the EU is a slow-moving and multi-faceted body that is notoriously bad at disseminating its decisions. What this means however, is that informing voters about the EU verdict during the survey experiment may well be the first opportunity that respondents have to actually consume this information. At the same time, given that the verdict is about to be issued within weeks, possibly within days of the election, we could collect voters' responses in the period between the revelation of the verdict and the actual election.

The research design we propose here is to imbed an experimental vignette in a survey conducted directly after the EUCJ verdict, but close to the election on PiS voters, that is voters who in 2015 cast their vote for PiS. There are roughly around 20% of the representative sample, which consists of all Poles eligible to vote. Although, roughly 40 % of respondents who are voters admit having voted for PiS, nearly half of the sample typically polled by CBOS are non-voters. For this reason, we are working with the pollsters on expanding the sample to target respondents who voted for PiS on 2015.

The survey will be conducted by the Center for Public Opinion Research (CBOS) in Warsaw and imbedded in the omnibus survey, which also means that we will be able to obtain a host of demographic variables that can be used to check the sample of treated and control PiS voters for balance on covariates that typically predict the PiS vote. Among them are questions about education, income, religiosity, urbanization of the area where one's household is located, and gender.

All interviews are conducted face-to-face. The question “Which party did you vote for in the 2015 parliamentary elections will be asked early in the survey, directly following the battery of demographic questions

Next, half of the respondent who indicated they voted for PiS will receive the following treatment: “On October 10, the European Union’s Court of Justice issued a verdict regarding the reforms of the Supreme Court and the National Council of the Judiciary, determining that both laws infringe upon European values. The bill was sent to the Tribunal following a vote in the European Council, which in an assembly of all cabinet ministers from all member states. The Court has obligated the legislature to make changes to the law, bringing it back to agreement with these values. Minister Ziobro, who was among the authors of the bill responded that the bill is an attack on Polish sovereignty. Do you believe that the ruling parliamentary majority will amend the bill in accordance with the EUCJ recommendations?”

Respondents can answer 1 “Very likely” 2 “quite likely” 3. “rather unlikely” 4. “very unlikely.”

The role of this question is not so much to solicit answers from the respondents (what they answer is of secondary interest) but to inform them of the verdict of the EUCJ. Because the verdict of the EUCJ is an outside and objective source of whether undermining the rule of law took place. Our expectation is that upon learning about it, PiS voters will update their beliefs that they are dealing with a closet autocrat.

Meanwhile, the control group of PiS voters will receive this question: “On August 4 1, the EU Commissioner for Health, Vytenisa Andriukaitis, released comprehensive statistics on the consumption of tobacco products, including e-cigarettes. Specifically, these data reveal that tobacco consumption by Europeans aged 18-24 has increased relative to previous years (it is currently at 29 %), predominantly due to the unregulated marketing of e-cigarettes. In response the EU has proposed to convene a special directives associated with labeling these new tobacco products and funding for tobacco use cessation programs in member states. Do you believe that EU engagement will help limit the numbers of young Europeans who smoke?”

Respondents can answer 1 “Very likely” 2 “quite likely” 3. “rather unlikely” 4. “very unlikely.”

The goal of this question is to provide voters with information that will leave them neutral with respect to the probability that the incumbent is a closet autocrat. As this part of our research is still in the pilot phase, we welcome all and any feedback on the research design and we are piloting alternative sources allowing voters to update their probability that the PiS incumbent is a closet autocrat. Specifically, we are considering the following:

- (1) United Nations Human Rights Council has initiated an investigation into a breach of rights of free expression of Polish judges facing disciplinary action for protesting the early retirement of members of the Supreme Court;
- (2) US State department has issued a statement criticizing the four recent judiciary reforms as undermining rule of law;
- (3) The International Council of Constitutional Lawyers has issued a chilling opinion criticizing the four recent judiciary reforms for ignoring basic principles of rule of law.

Voters who indicated they would not be voting or who indicated they would not be voting for PiS would only be asked the control question. The idea of asking the control question to those voters would be to verify whether indeed it is neutral.

Finally, at the end of the omnibus, voters would be asked the question: “If the election were to take place this coming Sunday, which party would you vote for?” The format of the question is identical to the one that was used to identify PiS voters in the first place.

Our expectation is that having updated the probability that the incumbent is a closet autocrat will lead former PiS voters to be less willing to cast their vote for PiS again.

There are several problems with this research design that we anticipate. The first is that if the EUCJ treatment proves to be ineffective in the pilot study, the three alternative treatments would have to rely on deceit.

The second potential problem is that there may be no light whatsoever between the EUCJ’s ruling and the date of the parliamentary elections. Having to run the survey after the elections would present a complication, because identifying PiS supporters would require voters being able to recall their choices from four years ago. In addition, the 2019 vote would already have been

cast. Yet if the timing of the survey and the election were to misalign in such a way, it could be used to corroborate another implication of the formal model: the occurrence of voter regret. Recall that one of the empirical implications of our formal model is that the citizen, following a vote cast in favor of the incumbent, will experience “rational regret” when that incumbent turns out to be a closet autocrat. Hence, based on the outcome of the pilot study and the timing of these two exogenous events—the parliamentary election date and the EUCJ ruling—we could field our survey following both events and test the expectation that after having the opportunity to update their beliefs regarding the type of incumbent they are facing, PiS voters will express regret about having voted for PiS. This could be operationalized as the answer to the question about voting intentions were the elections to take place the following Sunday.

## 5. CONCLUSION

In this paper, we have addressed the increasingly popular phenomenon of authoritarian back-sliding in recently consolidated democracies. While some scholars attribute it to polarization of the electorate (Svolik 2018), others focus on the polarization of elites stressing simultaneously the role of majoritarian political institutions that have the ability of shutting out opposition to the government. In this paper, we try to reconcile these two viewpoints. We also aspire to account for the possibility that voters support closet autocrats unknowingly as such closet autocrats reveal their authoritarian intentions gradually. By the time voters discover that whom they have elected into office is in fact a dictator, it is often too late. For an illustration, one need look no further than the Hungarian election of April 2018, followed by protests of tens of thousands in Budapest.

<sup>23</sup> This most recent Hungarian election stands out in particular because for the first time, the incumbent has secured a  $\frac{2}{3}$  majority in the legislature (allowing to amend the constitution) with only a plurality of the vote. In order to account for the gradual descent into authoritarianism observed in Hungary, but also possibly in Poland, the source of our analytic narrative, we propose a two-period signaling model, in which the incumbent can be one of two types. The first type has no authoritarian tendencies, but is simply ideologically to the right. The second type is a closet

---

<sup>23</sup> The Guardian “Thousands rally against Viktor Orban’s election victory in Budapest”, Sunday, April 2018

autocrat. The incumbent first chooses the level of reforms on an institutional dimension and then stands for reelection. Only upon being reelected does he get a chance to implement policy on the authoritarian-democratic dimension. This means that when the electorate is deciding whether or not to reelect the incumbent, there is uncertainty about his type. What our model allows us to account for, is the behavior of voters in cases where it is hard to resolve whether incumbents are parties of ideological conservatives or a party of closet autocrats.

#### A FEW MORE SENTENCES ABOUT WHAT WE FIND AND WHAT THIS MEANS

## APPENDIX A. EMPIRICAL APPENDIX

The four questions that measure the respondent's sensitivity to the authoritarian dimension asked the respondent to what extent he or she agrees with the following four separate statements.

- Democracy is superior to any other form of rule (*Authoritarian1*)
- For people like me, it does not matter whether the regime is authoritarian or democratic (*Indifferent*)
- Sometimes Non-democratic rule is better than democratic rule (*Authoritarian2*)
- Government by a strong leader is decidedly better than democratic rule (*StrongLeader*)

Respondents could “agree strongly”, “agree somewhat”, “rather disagree”, “strongly disagree” with the above statements. Higher values of these variables represent stronger disagreement. These four questions are tapping into something very different from just general attitudes to the opposition.

TABLE 2. Correlation matrix

	Polarization in Electorate	Indifferent	Authoritarian 2	Strong Leader	Authoritarian
Polarization in Electorate	1.00				
Indifferent	0.02 (0.50)	1.00			
Authoritarian2	0.11 (0.00)	0.42 (0.00)	1.00		
Strong Leader	0.19 (0.00)	0.38 (0.00)	0.48 (0.00)	1.00	
Authoritarian	-0.03 (0.35)	-0.18 (0.00)	-0.27 (0.00)	-0.17 (0.00)	1.00

Note: P-values in parentheses

## REFERENCES

- Ahlquist, John S, Nahomi Ichino, Jason Wittenberg and Daniel Ziblatt. 2018. “How Do Voters Perceive Changes to the Rules of the Game? Evidence from the 2014 Hungarian Elections.” *Journal of Comparative Economics* .
- Bates, Robert H. 2007. From case studies to social science: A strategy for political research. In *The Oxford Handbook of Comparative Politics*.
- Bermeo, Nancy. 2016. “On democratic backsliding.” *Journal of Democracy* 27(1):5–19.
- Gailmard, Sean and John W Patty. 2012. *Learning while governing: Expertise and accountability in the executive branch*. University of Chicago Press.
- Goemans, Hein E and Mark Fey. 2009. “Risky but rational: War as an institutionally induced gamble.” *The Journal of Politics* 71(1):35–54.
- Jenne, Erin K and Cas Mudde. 2012. “Can outsiders help?” *Journal of Democracy* 23(3):147–155.
- Levi, Margaret. 2004. “10 An analytic narrative approach to puzzles and problems.” *Problems and Methods in the Study of Politics* p. 201.
- Levitsky, Steven and Daniel Ziblatt. 2018. *How democracies die*. Crown.
- Lex Super Omnia Association of Prosecutors to the Prosecutor General*. 2019.
- Lorentzen, Peter. 2014. “China’s strategic censorship.” *American Journal of Political Science* 58(2):402–414.
- Lorentzen, Peter L, M Taylor Fravel and Jack Paine. 2015. “Using process tracing to evaluate formal models.”.
- Lust, Ellen and David Waldner. 2015. “Unwelcome Change: Understanding, Evaluating, and Extending Theories of Democratic Backsliding.” *US Agency for International Development* 11.
- Nalepa, Monika. 2010. *Skeletons in the closet: Transitional justice in post-communist Europe*. Cambridge University Press.
- Nalepa, Monika. 2016. “Party institutionalization and legislative organization: The evolution of agenda power in the polish parliament.” *Comparative Politics* 48(3):353–372.

of Justice of the European Union, Court. 2019a. “The Polish legislation concerning the lowering of the retirement age of judges of the Supreme Court is contrary to EU law.” Resolution(2002)3 adopted by the Committee of Ministers.

of Justice of the European Union, Court. 2019b. “The Polish legislation concerning the lowering of the retirement age of judges of the Supreme Court is contrary to EU law.”.

*Revised Statute of the European Commission for Democracy through Law.* 2002. Resolution(2002)3 adopted by the Committee of Ministers.

Sedelmeier, Ulrich. 2014. “Anchoring democracy from above? The European Union and democratic backsliding in Hungary and Romania after accession.” *JCMS: Journal of Common Market Studies* 52(1):105–121.

*Sędziowie pod Presją [Judges under Pressure].* 2019.

Serra, Gilles. 2012. “The Risk of Partyarchy and Democratic Backsliding Mexico’s 2007 Electoral Reform.” *Taiwan Journal of Democracy* 8(1).

Svolik, Milan. 2018. “When Polarization Trumps Civic Virtue: Partisan Conflict and the Subversion of Democracy by Incumbents.” *Unpublished Manuscript* .

Vanberg, Georg. 2000. “Establishing judicial independence in West Germany: The impact of opinion leadership and the separation of powers.” *Comparative Politics* pp. 333–353.