

# Formal Appendix

In the model presented in the main text, we assume that the ideologue's ideal point is fixed at  $X_I = (1, 0)$ . A more general version allows the ideologue's ideal point to vary in the policy space,  $X_I = (p_I, 0)$ , with  $p_I \in [0, 1]$ . In this section we outline this general version. The baseline model is a subcase, and will be presented in the following section.

## 1. THE EXTENDED MODEL

Recall the 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 authoritarian dimension and the policy dimension.
- (5) The game ends, and payoffs are collected.

**1.1. The citizen's reelection decision.** Consider the reelection decision facing the citizen. Suppose the incumbent has chosen reform  $i$  in the first period, and let  $\gamma$  denote the citizen's updated belief that the incumbent is a closet autocrat. Note that the citizen knows that if she reelects, the second period policy will either be  $(p, a) = (\min[p_i, i], 0)$  (if the incumbent is an ideologue) or  $(p, a) = (i, i)$  (if the incumbent is a closet authoritarian). If she chooses the opposition instead,

policy will be  $p = (0, 0)$ . Moreover, note that for the ideologue, setting  $i > p_I$  is always weakly dominated. Thus, the expected utilities of the options confronting a citizen are given by:

$$EU_C(\text{Reelect}) = -(1 - \gamma)(i - x_C)^2 - \gamma((i - x_C)^2 + \beta(i)^2)$$

$$EU_C(\text{Replace}) = -(x_C)^2$$

The citizen (weakly) prefers to reelect the incumbent as long as the following condition holds:

$$i \leq \frac{2x_C}{1 + \gamma\beta}$$

Consider what this threshold implies, depending on the beliefs that the citizen holds:

- If the citizen believes that the incumbent is a closet autocrat ( $\gamma = 1$ ), she will only reelect if reform is below the following threshold:

$$(1) \quad T_{CA} = \frac{2x_C}{\beta + 1}$$

- If the citizen believes that the incumbent is an ideologue ( $\gamma = 0$ ), she will only reelect if reform is below the following threshold:

$$(2) \quad T_I = 2x_C$$

- If the citizen's ex ante belief has not changed ( $\gamma = \alpha$ ), she will only reelect if reform is below the following threshold:

$$(3) \quad T_\alpha = \frac{2x_C}{\alpha\beta + 1}$$

## 1.2. The Incumbent's first-period choice.

1.2.1. *The Ideologue.* Consider the ideological government's reform choice in the first period. Since the reform decision has no direct payoff implications, the incumbent can make a reform choice that will ensure that it is reelected (and can implement a policy in period 2), or it can make a reform choice that ensures that it is removed from office and the opposition takes over in period 2. We assume that if the ideological incumbent is going to make a reform decision that will result in reelection, it will always set policy in the second period to  $(p, a) = (i, 0)$  – there is no reason

to adopt a reform that is more extreme than the policy the government will adopt in period 2. Because the ideologue will never move in the authoritarian direction, we can simplify the utility functions and focus only on the ideological dimension. Payoffs are given by:

$$(4) \quad EU_I(\text{Reelect}) = -(i - p_I)^2 - \theta(i - p_I)^2 + b$$

$$(5) \quad EU_I(\sim \text{Reelect}) = -p_I^2$$

For the ideologue, the payoff to making an institutional proposal that will get him reelected requires that (4) is weakly greater than (5). This, in turn, requires that the cutoff  $i$  that leads to reelection is greater than this threshold, i.e., the ideologue is willing to set  $i$  above the following threshold in order to be reelected, but is not willing to set it at or below:

$$(6) \quad R_I = p_I - \frac{\sqrt{b + p_I^2}}{\sqrt{1 + \theta}}$$

1.2.2. *The Closet Autocrat.* Consider the closet autocrat's reform choice in the first period. It can make a reform choice that will ensure that it is reelected (and can implement a policy in period 2), or it can make a reform choice that ensures that it is removed from office and the opposition takes over in period 2. In order to be reelected, the closet autocrat must choose an institutional reform that is sufficiently moderate to make the citizen willing to reelect. We assume that if the autocrat is going to choose a reform that results in reelection, it will always set policy in the second period to  $(p, a) = (i, i)$  – there is no reason to adopt a reform that is more extreme than the policy the government will adopt in period 2. Payoffs are given by:

$$(7) \quad EU_{CA}(\text{Reelect}) = -(2(i - 1)^2) - \theta(2(i - 1)^2) + b$$

$$(8) \quad EU_{CA}(\sim \text{Reelect}) = -2$$

For the closet autocrat, the payoff to making an institutional proposal that will get him reelected requires that (7) is weakly greater than (8). This, in turn, requires that the cutoff  $i$  that leads to reelection is greater than the following threshold, i.e., the autocrat is willing to set  $i$  above the following threshold in order to be reelected, but is not willing to set it at or below:

$$(9) \quad R_{CA} = 1 - \frac{\sqrt{2+b}}{\sqrt{2(1+\theta)}}$$

1.2.3. *Ordering of thresholds.* Algebraic manipulation shows that  $R_I < R_{CA}$ . Moreover, it is immediate that  $T_I > T_\alpha > T_{CA}$ .

1.2.4. *Off-equilibrium path beliefs.* The off-equilibrium path beliefs we specify are loosely based on the logic of the intuitive criterion and the assumption that beliefs are monotonic in  $i$ . Note that the closet autocrat is not willing to propose  $i \leq R_{CA}$  and the ideologue is not willing to propose  $i \leq R_I$ . There are several possibilities we must consider:

**Case 1:**  $R_{CA} > p_I > R_I$ :

- (1) For any proposal  $i > R_{CA}$ , the belief is that the proposal comes from the closet autocrat, i.e.,  $\gamma = 1$ . This is intuitive since the ideologue never has an incentive to make a proposal above his ideal point, and the closet autocrat is willing to propose above  $R_{CA}$ .
- (2) For any proposal  $i \leq R_{CA}$ , the belief is that the proposal comes from the ideologue. Above  $R_I$ , the ideologue is willing to make such a proposal. And below  $R_I$ , the cost of this proposal is less for the ideologue than the CA. So set  $\gamma = 0$ .

**Case 2:**  $p_I > R_{CA} > R_I$ :

- (1) For any proposal  $i > p_I$ , the belief is that the proposal comes from the closet autocrat, i.e.,  $\gamma = 1$ . This is intuitive since the ideologue never has an incentive to make a proposal above his ideal point.
- (2) For any proposal  $i \in (R_{CA}, p_I]$ , the belief is that the proposal could come from either, i.e.,  $\gamma = \alpha$ .

- (3) For any proposal  $i \leq R_{CA}$ , neither incumbent is expected to make this proposal, but the cost of this mistake is less for the ideologue. So set  $\gamma = 0$ .

**1.3. Equilibria.** Depending on the relative position of  $p_I$  and the thresholds imposed by the citizen – i.e., the level of citizen and elites’ polarization, and uncertainty on the type of incumbent – we have seven different scenarios, leading to just as many types of equilibria, described below.

For all the equilibria that follow, note that the citizen will reelect the incumbent despite *knowing for certain* he is the closet autocrat only if the latter proposes a reform  $i \leq T_{CA}$ . If she observes a proposal between  $T_{CA}$  and  $T_\alpha$ , she will reelect only if she is uncertain whether she is facing the ideologue or the closet autocrat. If the incumbent proposes reforms above  $T_\alpha$  and below or at  $T_I$ , she will reelect only if she is certain the proposal comes from the ideologue. Above  $T_I$ , the citizen will replace the incumbent with the opposition.

**Case 1: No Type is Reelected.** *Claim. If and only if  $T_I \leq R_I$ , the ideologue will propose  $i = p_I$ , the closet autocrat  $i = x_{CA}$ , and no type of incumbent is reelected. The equilibrium is unique.*

( $\Rightarrow$ ) Suppose not: the ideologue could set  $i \leq T_I$  to be reelected. But since  $T_I \leq R_I$ , the ideologue would rather propose its ideal point and be replaced. As shown above,  $R_I < R_{CA}$ , so the reasoning is similar for the closet autocrat. So it has to be that  $i > R_I$  for the ideologue and  $i > R_{CA}$  for the closet autocrat, and each will set its ideal point.<sup>1</sup>

( $\Leftarrow$ ) Suppose  $i = p_I$  for the ideologue,  $i = x_{CA} = 1$  for the closet autocrat, no type is reelected, and  $R_I < T_I$ . There are two cases: either  $p_I \leq T_I$ , or  $p_I > T_I$ . If  $p_I \leq T_I$ , if  $i_{CA} = 1$ , then if the ideologue sets  $i = p_I$ , it will get reelected, a contradiction. If  $p_I > T_I$ , but  $R_I < T_I$ , it means there exists a set of policies that the ideologue would pick in order to get reelected, which it prefers to setting  $i = p_I$  and not getting reelected, which is also a contradiction. So  $T_I \leq R_I$  is a necessary

<sup>1</sup> The incumbent should be indifferent between any proposal that will ensure reelection – the payoff is always 0. From now on we will assume that in such scenario the incumbent will propose its ideal point, given a small chance that the citizen may mistakenly reelect him.

condition for the equilibrium described above.

**Case 2: Both Types are Reelected with a Separating Equilibrium.** *Claim. If and only if  $R_I < T_I$ ,  $p_I \leq T_{CA}$ , and  $R_{CA} < T_{CA}$ , the ideologue will propose  $i = p_I$ , and the closet autocrat  $i = T_{CA}$ . Both types of incumbent will be reelected. The equilibrium is unique.*

( $\Rightarrow$ )  $T_{CA}$  is the threshold under which the citizen is willing to reelect the incumbent, if she is sure it is the closet autocrat. The citizen is therefore always willing to reelect if she observes  $i \leq T_{CA}$ . If the ideologue sets  $i = p_I$ , it knows it will be reelected and maximizes its utility, so it has no reason to deviate from its ideal point. The closet autocrat can either set  $i \leq T_{CA}$  and be reelected, or set its ideal point and not be reelected. But by definition of  $R_{CA}$ , the closet autocrat prefers to set any point  $i > R_{CA}$  and be reelected rather than setting  $i = x_{CA}$  and be replaced. So it will set  $i$  equal to the point closest to its ideal point (furthest to the right) that will get it reelected, i.e.,  $i = T_{CA}$ . For any point  $i < T_{CA}$ ,  $u_{CA}(i) < u_{CA}(T_{CA})$ , since it is further from its ideal point. With any point  $i > T_{CA}$ , it will not be reelected: the citizen will know that the proposed policy comes from the closet autocrat, since  $p_I \leq T_{CA}$ , and by definition of  $T_{CA}$ , it is not willing to reelect the closet autocrat proposing any policy above it. And because  $R_{CA} < T_{CA}$ ,  $u_{CA}(i) < u_{CA}(T_{CA})$  for any  $i > T_{CA}$ .

( $\Leftarrow$ ) Suppose the ideologue proposed  $i = p_I$ , the closet autocrat  $T_{CA}$ , both are reelected, but any of the following does not hold:  $R_I < T_I$ ,  $p_I \leq T_{CA}$ ,  $R_{CA} < T_{CA}$ .

If  $R_{CA} \geq T_{CA}$ , the closet autocrat is not willing to set  $i = T_{CA}$ , a contradiction. If  $R_I \geq T_I$ , then  $R_{CA} > T_{CA}$ , since  $R_I < R_{CA}$ , again a contradiction. If  $T_{CA} < p_I$ , then the closet autocrat may want to pool at  $p_I$  (a point closer to its ideal point), so it would not be a separating equilibrium.

**Case 3.1: Only the Ideologue is Reelected at  $p_I$ .** *Claim. If and only if  $R_I < T_I$ ,  $p_I \leq R_{CA}$ ,  $T_{CA} \leq R_{CA}$ , and  $p_I \leq T_I$ , the ideologue will set  $i = p_I$ , and the closet autocrat will set  $i = x_{CA}$ .*

*The ideologue will be reelected, the closet autocrat will not. The equilibrium is unique.*

( $\Rightarrow$ ) Since  $p_I \leq R_{CA}$ , the closet autocrat is not willing to pool at  $p_I$ . If the citizen observes  $i = p_I$ , she will know with certainty that it was the ideologue to propose it. And because  $p_I \leq T_I$ , the citizen is willing to reelect the ideologue. Setting its ideal point, knowing it will be reelected, is the best possible outcome for the ideologue, and it will do so.

The closet autocrat, as above, is not willing to pool at  $p_I$ . And because  $T_{CA} \leq R_{CA}$ , there is no policy the closet autocrat is willing to set for which the citizen will reelect it, even knowing that it is the closet autocrat who proposed it. The closet autocrat is hence better off by setting its ideal point, and not be reelected.

( $\Leftarrow$ ) Suppose the ideologue picks  $i = p_I$  and gets reelected, the closet autocrat picks  $x_{CA}$  and does not, and any of the following does not hold:  $R_I < T_I$ ,  $p_I \leq R_{CA}$ ,  $T_{CA} \leq R_{CA}$ , or  $p_I \leq T_I$ .

If  $p_I > T_I$ , then the ideologue will not get reelected, a contradiction. If  $R_I \geq T_I$ ,  $p_I$  also has to be greater than  $T_I$ , the same contradiction.

If  $p_I > R_{CA}$ , the closet autocrat would be willing to pool at  $i = p_I$ , and this would be not an equilibrium. If  $T_{CA} > R_{CA}$ , the closet autocrat could get reelected by picking  $T_{CA}$ , again a contradiction.

**Case 3.2: Only the Ideologue is Reelected at  $T_I$ .** *Claim. If and only if  $R_I < T_I$ ,  $T_I \leq R_{CA}$ ,  $T_I < p_I$ , the ideologue will set  $i = T_I$ , and the closet autocrat will set  $i = x_{CA}$ . The ideologue will be reelected, the closet autocrat will not. The equilibrium is unique.*

( $\Rightarrow$ ) The closet autocrat is not willing to propose any policy for which the citizen would reelect; and is better off setting its ideal point,  $x_{CA} = 1$ , knowing it will not be reelected.

If the ideologue sets its ideal point, it will not be reelected, as  $p_I > T_I$ . The closest point to its preferred policy which will still ensure reelection is  $T_I$ . Since  $R_I < T_I$ , the ideologue prefers setting

$T_I$  and get reelected than setting its ideal point and being replaced.

The citizen knows the closet autocrat would not set  $T_I$  as its proposed policy; observing  $T_I$ , she knows it can only be the ideologue proposing it and, by definition, she will reelect the ideologue.

( $\Leftarrow$ ) Suppose the ideologue picks  $i = T_I$  and gets reelected, while the closet autocrat sets  $i = 1$  and does not, and suppose any of the conditions does not hold:  $R_I < T_I$ ,  $T_I \leq R_{CA}$ ,  $T_I < p_I$ .

If  $R_I \geq T_I$ , the ideologue is not willing to pick  $T_I$ , a contradiction. If  $T_I \geq p_I$ , the ideologue would be better off by setting  $i = p_I$ , so the one described above would not be an equilibrium. If  $T_I > R_{CA}$ , the closet autocrat may want to mimic the ideologue by picking  $T_I$  as well; in which case, no type would get reelected, as the citizen would only reelect after observing  $T_I$  if she is certain the incumbent is the ideologue.

**Case 3.3: Only the Ideologue is Reelected at  $R_{CA}$ .** *Claim. If and only if  $R_I < T_I$ ,  $R_{CA} < p_I$ ,  $T_\alpha \leq R_{CA} \leq T_I$ , the citizen will reelect if  $i \leq T_I$ , the ideologue will set  $i = R_{CA}$ , and the closet autocrat will set  $i = x_{CA}$ . The ideologue will be reelected, the closet autocrat will not. The equilibrium is unique.*

( $\Rightarrow$ ) Since  $R_{CA} < p_I$ , the ideologue is aware that the closet autocrat may mimic its proposed policy, if it decides to set it at its ideal point,  $p_I$ . And because both  $R_{CA}$  and  $p_I$  are above  $T_\alpha$ , the threshold under which proposing the same policy would get them reelected, setting any policy above  $R_{CA}$  would mean replacement for the ideologue.  $R_I < T_I$  implies that there exists a set of policies that the ideologue would pick to ensure reelection.  $R_I < R_{CA}$  means that proposing  $R_{CA}$  and getting reelected is preferred to setting  $p_I$  and not getting reelected. By definition,  $R_{CA}$  is the point furthest to the right – in this case, also closest to  $p_I$  – the closet autocrat is not willing to propose. In sum, the ideologue would rather set  $i = R_{CA}$ , which ensures reelection, than  $i = p_I$ , which would not get it reelected. The ideologue thus clearly signals its type by proposing a policy the closet autocrat never would. Since  $R_{CA} \leq T_I$ , and the citizen observing  $R_{CA}$  is sure of the



incumbent's type, the ideologue will get reelected.

The closet autocrat is, by definition, not willing to pool at  $R_{CA}$ . Any policy proposals above, however, would not ensure reelection, since  $R_{CA} > T_{CA}$ . The closet autocrat is hence better off setting its ideal point  $x_{CA} = 1$ , as it will not get reelected, no matter what.

( $\Leftarrow$ ) Suppose the ideologue picks  $i = R_{CA}$  and gets reelected, the closet autocrat picks  $x_{CA}$ , and does not, and suppose any of the following does not hold:  $R_I < T_I$ ,  $R_{CA} < p_I$ ,  $T_\alpha \leq R_{CA} \leq T_I$ .

If  $R_{CA} > T_I$ , if the ideologue picks  $R_{CA}$ , it will not get reelected, a contradiction. If  $R_{CA} < T_\alpha$ , and  $R_{CA} < p_I$ , the ideologue would rather propose its ideal point, the closet autocrat would also rather pool at  $i = p_I$ , and they could both get reelected: the choices above would not lead to an equilibrium.

If  $R_I > T_I$ ,  $R_{CA}$  would also be above  $T_I$ , and the first contradiction found would apply.

If  $R_{CA} \geq p_I$ , the ideologue could pick  $p_I$  without the risk of the closet autocrat pooling, so picking  $R_{CA}$  would not be optimal for the ideologue.

**Case 4.1: Both Types are Reelected at  $T_\alpha$ .** *Claim. If and only if  $R_I < T_I$ ,  $R_{CA} < p_I$ ,  $R_{CA} < T_\alpha \leq p_I$ , both the ideologue and the closet autocrat will set  $i = T_\alpha$ , and will both get reelected.*

( $\Rightarrow$ ) Because  $R_{CA} < p_I$ , the ideologue knows the closet autocrat may mimic its behavior if it sets its ideal point. And because  $p_I \geq T_\alpha$ , if both set  $i = p_I$ , none of them will get reelected. However, because  $R_I < R_{CA} < T_\alpha$ , there exists a set of policies that both types would set in order to get reelected, even if they differ from their ideal point.  $T_\alpha$  is the closest point to both the ideologue and the closet autocrat's ideal points that will get them reelected, hence they will both set  $i = T_\alpha$ .

Note that neither has incentives to deviate. If the ideologue deviates by setting any point closer or equal to its ideal point, the closet autocrat may mimic it, but because those points would be

above  $T_\alpha$ , neither would get reelected. Such points would be also above  $R_I$  and  $R_{CA}$ , they would both prefer a policy that gets them reelected, so they would be better off setting  $T_\alpha$ .

If the ideologue were to deviate by setting any point below  $T_\alpha$ , it would still get reelected, but with a policy further away from its own ideal point. Again, it would rather set  $i = T_\alpha$ . A similar reasoning applies for the closet autocrat.

( $\Leftarrow$ ) Suppose both types of incumbent choose  $T_\alpha$  and get reelected, but either  $R_I \geq T_I$ ,  $R_{CA} \geq p_I$ ,  $R_{CA} \geq T_\alpha$ , or  $p_I < T_\alpha$ .

If  $p_I < T_\alpha$ , the ideologue would rather pick  $p_I$ , so the choices above would not be an equilibrium.

If  $R_{CA} \geq T_\alpha$ , the closet autocrat is not willing to pick  $T_\alpha$  to get reelected, a contradiction. If  $R_{CA} \geq p_I$ , then  $R_{CA} \geq T_\alpha$ , and the same reasoning applies. If  $R_I \geq T_I$ , neither type of incumbent is willing to pick  $T_\alpha$ , a contradiction.

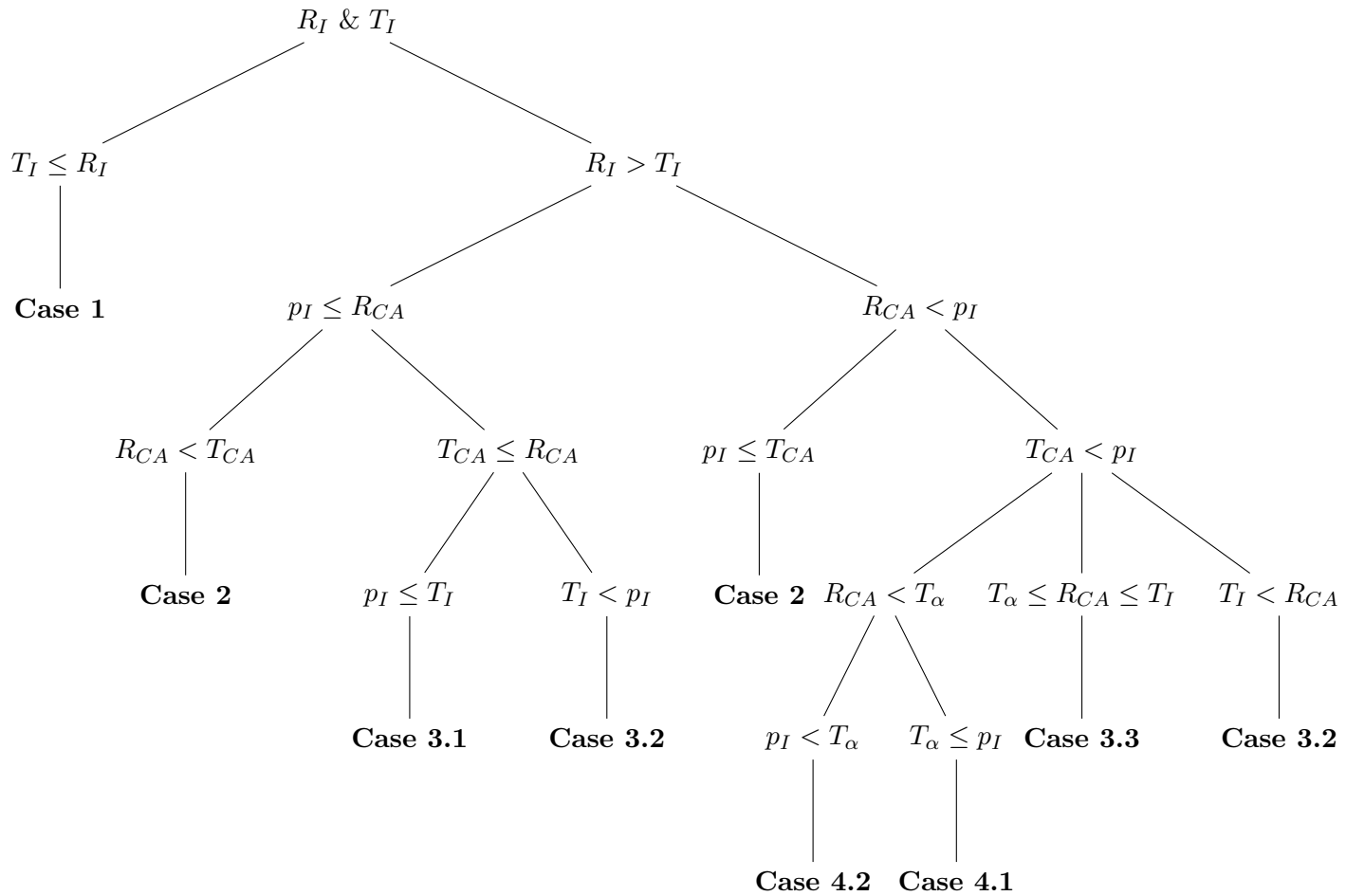
**Case 4.2: Both Types are Reelected at  $i = p_I$ .** *Claim. If and only if  $R_I < T_I$ ,  $R_{CA} < p_I$ ,  $T_{CA} < p_I < T_\alpha$ , both the ideologue and the closet autocrat will set  $i = p_I$ , and will both get reelected.*

( $\Rightarrow$ ) Since  $p_I < T_\alpha$ , the ideologue knows that, by setting its ideal point, it will get reelected no matter what the closet autocrat proposed. It has hence no incentives to deviate from  $p_I$ .

The closet autocrat prefers setting  $p_I$  and getting reelected than its ideal point and not getting reelected, since  $R_{CA} < p_I$ . Note that it also has no incentives to deviate. Another point that would ensure reelection would be  $T_{CA}$ , but it is better off setting  $p_I$  as the latter is closer to its ideal point. Any other point above  $p_I$ , since they would never be picked by the ideologue, would reveal its type as closet autocrat, and given that  $p_I > T_{CA}$ , it would not get reelected. Given  $R_{CA} < p_I$ , it would prefer  $p_I$  and reelection to a higher point and no reelection. So the closet autocrat's optimal strategy is also to propose  $p_I$ .

( $\Leftarrow$ ) Suppose both the ideologue and the closet autocrat set  $i = p_I$  and get reelected, but any of the following does not hold:  $R_I < T_I$ ,  $R_{CA} < p_I$ ,  $T_{CA} < p_I < T_\alpha$ .

If  $p_I \leq R_{CA}$ , the closet autocrat is not willing to set  $i = p_I$ , a contradiction. If  $p_I \leq T_{CA}$ , the closet autocrat would rather set  $T_{CA}$  instead of  $p_I$ , to get closer to its ideal point, and the choices above would not be an equilibrium. If  $p_I \geq T_\alpha$ , the citizen observing  $p_I$  would not be willing to reelect under the uncertainty on the type of incumbent she is facing. If  $R_I \geq T_I$ , it has to be that  $p_I \geq T_\alpha$ , and the the same logic applies.



1.4. **Completeness of Equilibria.** Note that the conditions above are exhaustive, so there are no other equilibria than the ones listed above. The tree above shows how the conditions from the

previous sections cover all possible scenarios.

## 2. THE BASELINE MODEL

We now consider our baseline model, which is one particular case of the model above, with  $p_I = 1$ . The closet autocrat and the ideologue share the same ideal point in the ordinary policy dimension.

**2.1. Reelection Decision and First-Period Choice.** The citizen's reelection decision is the same as above, so that the thresholds  $T_{CA}$ ,  $T_\alpha$ ,  $T_I$  are defined in (1), (2), (3).

Similarly, the closet autocrat's first-period choice will be determined by the threshold  $R_{CA}$  defined in (9).

The ideologue's payoffs for setting reform  $i$  and getting reelected are:

$$EU_I(\text{Reelect}) = -(i - 1)^2 - \theta(i - 1)^2 + b$$

and in case of no reelection:

$$EU_I(\sim \text{Reelect}) = -1$$

It follows that the ideologue is willing to set  $i$  at or above the following threshold in order to be reelected, but is not willing to set it below:

$$(10) \quad R_I = 1 - \frac{\sqrt{b+1}}{\sqrt{1+\theta}}$$

**2.2. Ordering of thresholds.** As above, it still holds that  $R_I < R_{CA}$  and that  $T_I > T_\alpha > T_{CA}$ .

**2.3. Off-equilibrium path beliefs.** We now only need to consider the case in which  $x_I > R_{CA} > R_I$ . We assume that:

- (1) For any proposal  $i \in (R_{CA}, 1]$ , the belief is that the proposal could come from either, i.e.,  $\gamma = \alpha$ .
- (2) For any proposal  $i \in (R_I, R_{CA}]$ , the belief is that only the ideologue would make the proposal, i.e.,  $\gamma = 0$ .

- (3) For any proposal  $i \leq R_I$ , neither incumbent is expected to make this proposal, but the cost of this mistake is less for the ideologue. So set  $\gamma = 0$ .

2.4. **Equilibria.** With  $p_I = 1$ , it will never be the case that  $p_I \leq R_{CA}$ . Case 2 and 3.1 are not possible in this setting, and we are left with the following five cases.

**Case 1 (Case 1 above): No Type is Reelected.** *Claim. If and only if  $T_I \leq R_I$ , the ideologue will propose  $i = p_I = 1$ , the closet autocrat  $i = x_{CA} = 1$ , and no type of incumbent is reelected. The equilibrium is unique.*

( $\Rightarrow$ ) Suppose not: the ideologue could set  $i \leq T_I$  to be reelected. But since  $T_I \leq R_I$ , the ideologue would rather propose its ideal point and be replaced. As shown above,  $R_I < R_{CA}$ , so the reasoning is similar for the closet autocrat. So it has to be that  $i > R_I$  for the ideologue and  $i > R_{CA}$  for the closet autocrat, and each will set his ideal point.

( $\Leftarrow$ ) Suppose  $i = p_I = 1$  for the ideologue,  $i = x_{CA} = 1$  for the closet autocrat, no type is reelected, and  $R_I > T_I$ . However, because  $R_I > T_I$ , there exists a set of policies that the ideologue would pick in order to get reelected, which it prefers to setting  $i = 1$  and not getting reelected, a contradiction. So  $T_I \leq R_I$  is a necessary condition for the equilibrium described above.

**Case 2.1 (Case 3.2 above): Only the Ideologue is Reelected at  $T_I$ .** *Claim. If and only if  $R_I < T_I \leq R_{CA}$ , the ideologue will set  $i = T_I$ , and the closet autocrat will set  $i = x_{CA}$ . The ideologue will be reelected, the closet autocrat will not. The equilibrium is unique.*

( $\Rightarrow$ ) The closet autocrat is not willing to propose any policy for which the citizen would reelect; and is better off setting its ideal point,  $x_{CA} = 1$ , knowing it will not be reelected.

If the ideologue sets its ideal point, it will not be reelected, since  $p_I > R_{CA} \geq T_I$ . The closest point to its preferred policy which will still ensure reelection is  $T_I$ . Since  $R_I < T_I$ , the ideologue prefers

setting  $T_I$  and get reelected than setting its ideal point and being replaced.

The citizen knows the closet autocrat would not set  $T_I$  as its proposed policy; observing  $T_I$ , she knows it can only be the ideologue proposing it and, by definition, she will reelect the ideologue.

( $\Leftarrow$ ) Suppose the ideologue picks  $i = T_I$  and gets reelected, while the closet autocrat sets  $i = 1$  and does not, and suppose one of the conditions does not hold:  $R_I < T_I$ , or  $T_I \leq R_{CA}$ .

If  $R_I \geq T_I$ , the ideologue is not willing to pick  $T_I$ , a contradiction. If  $T_I > R_{CA}$ , the closet autocrat may want to mimic the ideologue by picking  $T_I$  as well; in that case, no type would get reelected, as the citizen would only reelect after observing  $T_I$  if she is certain the incumbent is the ideologue.

**Case 2.2 (Case 3.3 above): Only the Ideologue is Reelected at  $R_{CA}$ .** *Claim. If and only if  $T_\alpha \leq R_{CA} \leq T_I$ , the citizen will reelect if  $i \leq T_I$ , the ideologue will set  $i = R_{CA}$ , and the closet autocrat will set  $i = x_{CA}$ . The ideologue will be reelected, the closet autocrat will not. The equilibrium is unique.*

( $\Rightarrow$ ) Since  $R_{CA} < p_I = 1$ , the ideologue is aware that the closet autocrat may mimic its proposed policy, if it decides to set it at its ideal point. And because both  $R_{CA}$  and  $p_I$  are above  $T_\alpha$  – the threshold under which proposing the same policy would get them reelected – setting any policy above  $R_{CA}$  would mean replacement for the ideologue.  $R_I < R_{CA} \leq T_I$  implies that there exists a set of policies that the ideologue would pick to ensure reelection.  $R_I < R_{CA}$  means that proposing  $R_{CA}$  and getting reelected is preferred to setting  $p_I$  and not getting reelected. By definition,  $R_{CA}$  is the point furthest to the right – also closest to  $p_I$  – the closet autocrat is not willing to propose. In sum, the ideologue would rather set  $R_{CA}$ , which ensures reelection, than  $p_I$ , which would not get it reelected. The ideologue thus clearly signals its type by proposing a policy the closet autocrat never would. Since  $R_{CA} \leq T_I$ , and the citizen observing  $R_{CA}$  is sure of the incumbent’s type, the ideologue will get reelected.

The closet autocrat is, by definition, not willing to pool at  $R_{CA}$ . Any policy proposals above,

however, would not ensure reelection, since  $R_{CA} > T_{CA}$ . The closet autocrat is hence better off setting its ideal point  $x_{CA} = 1$ , as it will not get reelected, no matter what.

( $\Leftarrow$ ) Suppose the ideologue picks  $i = R_{CA}$  and gets reelected, the closet autocrat picks  $x_{CA}$ , and does not, and suppose one of the following does not hold:  $T_\alpha \leq R_{CA}$ , or  $R_{CA} \leq T_I$ .

If  $R_{CA} > T_I$ , then if the ideologue picks  $R_{CA}$ , it will not get reelected, a contradiction. If  $R_{CA} < T_\alpha$ , the ideologue would rather propose  $T_\alpha$ , a policy closer to its ideal point that would still ensure reelection. The closet autocrat would also rather pool at  $T_\alpha$ , and they could both get reelected: the choices above would not lead to an equilibrium.

**Case 3.1 (Case 4.1 above): Both Types are Reelected at  $T_\alpha$ .** *Claim. If and only if  $R_{CA} < T_\alpha < 1$ , both the ideologue and the closet autocrat will set  $i = T_\alpha$ , and will both get reelected.*

( $\Rightarrow$ ) Because  $R_{CA} < p_I$ , the ideologue knows the closet autocrat may mimic its behavior if it sets its ideal point. And because  $p_I \geq T_\alpha$ , if both set  $i = p_I$ , none of them will get reelected. However, because  $R_I < R_{CA} < T_\alpha$ , there exists a set of policies that both types would set in order to get reelected, even if they differ from their ideal point.  $T_\alpha$  is the closest point to both the ideologue and the closet autocrat's ideal points that will get them reelected, hence they will both set  $i = T_\alpha$ .

Note that neither has incentives to deviate. If the ideologue deviates by setting any point closer or equal to its ideal point, the closet autocrat may mimic it, but because those points would be above  $T_\alpha$ , neither would get reelected. Such points would be also above  $R_I$  and  $R_{CA}$ , they would both prefer a policy that gets them reelected, so they would be better off by setting  $T_\alpha$ .

If the ideologue were to deviate by setting any point below  $T_\alpha$ , it would still get reelected, but with a policy further away from its own ideal point. It would rather set  $i = T_\alpha$ .

A similar reasoning applies for the closet autocrat.

( $\Leftarrow$ ) Suppose both types of incumbent choose  $T_\alpha$  and get reelected, but either  $R_{CA} \geq T_\alpha$ , or  $p_I \leq T_\alpha$ .

If  $p_I \leq T_\alpha$ , the ideologue would rather pick  $p_I$ , and  $T_\alpha$  could not be an equilibrium proposal. If  $R_{CA} \geq T_\alpha$ , the closet autocrat is not willing to pick  $T_\alpha$  to get reelected, a contradiction.

**Case 3.2 (Case 4.2 above): Both Types are Reelected at 1.** *Claim. If and only if  $T_\alpha \geq 1$ , both the ideologue and the closet autocrat will set  $i = 1$ , and will both get reelected.*

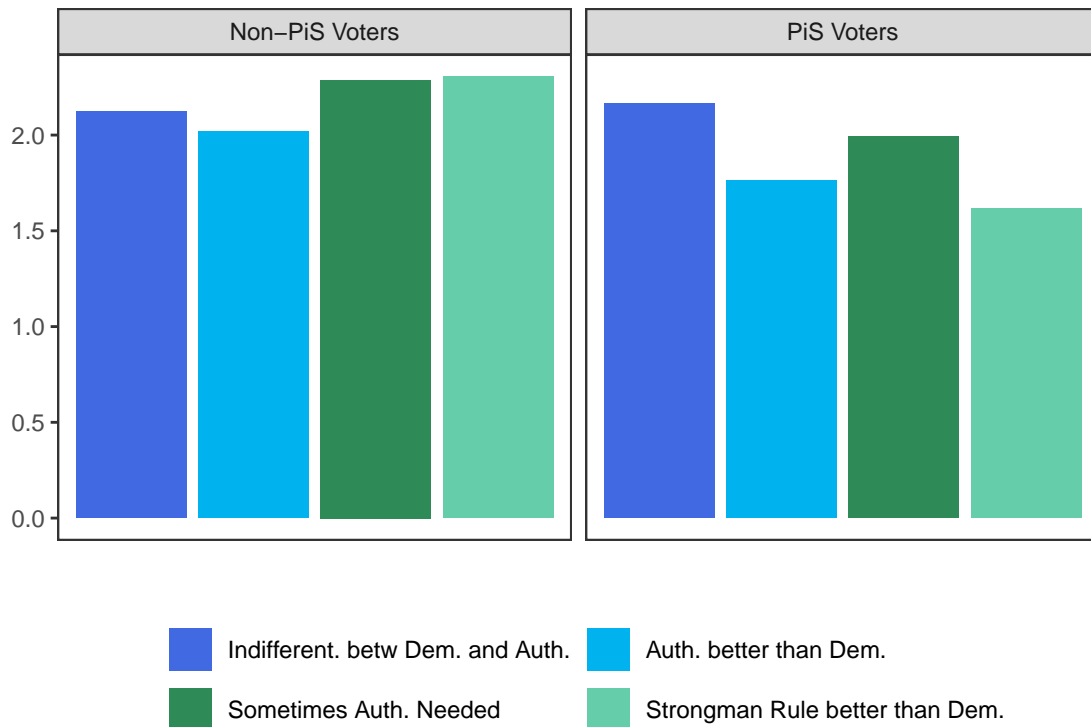
( $\Rightarrow$ ) Both the ideologue and the closet autocrat propose their ideal point and get reelected, so they have no incentive to deviate. Because  $T_\alpha$  is above the policy they both propose, the citizen is willing to face the uncertainty of the incumbent's type, and will reelect the incumbent if she observes  $i = 1$ .

( $\Leftarrow$ ) Suppose both the ideologue and the closet autocrat set  $i = 1$  and get reelected, but  $T_\alpha < 1$ . This is a contradiction: if they both propose a policy above  $T_\alpha$ , the citizen will still be uncertain about the incumbent's type, and is not willing to reelect him.



# Empirical Appendix

## EXPRESSED PREFERENCES FOR DEMOCRACY



**Figure A1.** *Differences in Authoritarian versus Democratic Values among PiS and non-PiS Polish voters. The answers were recoded to increase with pro-democratic values.*

Figure A1 provides survey evidence that supporters of the ruling Law and Justice party (PiS) report to value democracy and the rule of law just as much as those who cast their votes for parties in the opposition. The data is from an August 2017 survey conducted by the Center for Public Opinion Research.<sup>1</sup> The four questions that measure the respondent's sensitivity to the

<sup>1</sup> The reference survey for figure A1 is the Omnibus CBOS Survey number 2017\_20. Our experiment was conducted within a different survey, in 2019.

authoritarian dimension asked the respondent to what extent he or she agrees with the following four separate statements.

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

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. The responses were recoded so that all four answers increase with pro-democratic values. 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.

#### SURVEY DESCRIPTION

We embedded our experiment in a survey conducted by the Center for Public Opinion Research (CBOS, Centrum Badania Opinii Społecznej). It was conducted between November 23 and November 27, 2019 through face-to-face interviews with 725 respondents sampled to be representative of the population of Polish residents aged 18 and over who voted in national elections to the Lower and Upper houses of the legislature in the previous month (October 13, 2019). Individual respondents were selected at random with layered stratification corresponding to the level of:

- (1) Województwo (one of Poland's 16 regions)
- (2) Size of city or village (one of 6 classes)

From 80 strata created in this way, the Universal Electronic System for Population Registration (PESEL) is used to sample individuals in proportion to their size, taking into account the level of sample implementation that is characteristic for them. The level of implementation is estimated on the basis of data from nationwide research conducted by CBOS during the year preceding the survey. The selection procedure requires the designation of three types of sampling units. The

first are territorial units defined in accordance with the standards of the Central Statistical Office, taking into account the share of rural and urban areas in each wojewodztwo, as well as with the distinction of city districts in Warsaw, Łódź, Poznań and Wrocław. The second tier sampling units correspond to individual streets, groups of adjacent streets or villages, depending on the type of territorial unit. Individuals from the population are tertiary sampling units.

Each layer is independently drawn. First degree sampling units are drawn without returns with probabilities proportional to their estimated size. From each territorial unit selected in the first step, exactly one second-degree drawing unit is drawn irreversibly, i.e. one small area, with a probability proportional to its size. The third stage of sampling involves the selection of exactly nine observation units by means of a simple non-returnable drawing from each selected second-degree unit. Upon completion of fieldwork, the resulting dataset underwent a rim weighting procedure, taking into account the following socio-demographic characteristics:

- Gender
- Age
- Size of the place of residence
- Wojewodztwo
- Education.

The overweight sample reflects the distribution of the above-mentioned socio-demographic characteristics observed in the studied population. The demographic data used in the sampling and weighing process of the data set come from the resources of the Central Statistical Office (demographic databases and reports from the study “Labor Force Survey”) and from the state PESEL register.

## DESCRIPTIVE STATISTICS

The following tables report the distribution of what for our purposes are the main questions of interest within the survey. We distinguish between the entire sample and PiS voters.

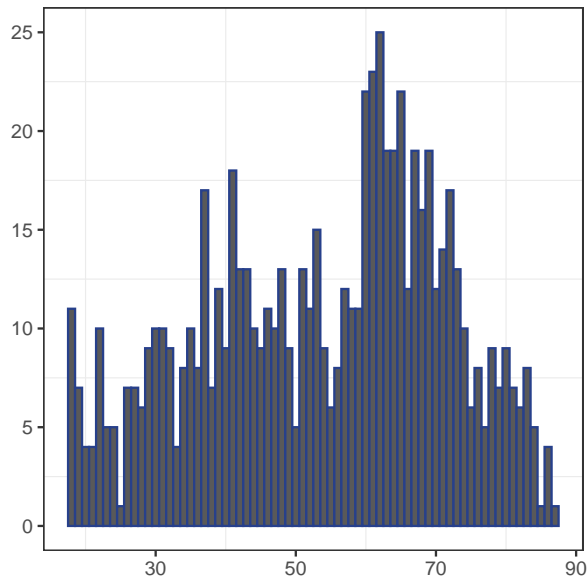
**Party Voted in the Last Election**

Party	Counts	Percentage
KW Polskie Stronnictwo Ludowe (PSL, Kukiz15)	50	6.90
KW Prawo i Sprawiedliwo (PiS, Porozumienie, SP)	356	49.10
KW Sojusz Lewicy Demokratycznej (SLD, Wiosna, Lewica Razem)	72	9.93
KW Konfederacja Wolno i Niepodlego	27	3.72
KKW Koalicja Obywatelska PO .N IPL Zieloni	173	23.86
Other	3	0.41
Vote not valid	2	0.28
Hard to say	17	2.34
Refuse to answer	25	3.45
Total	725	100

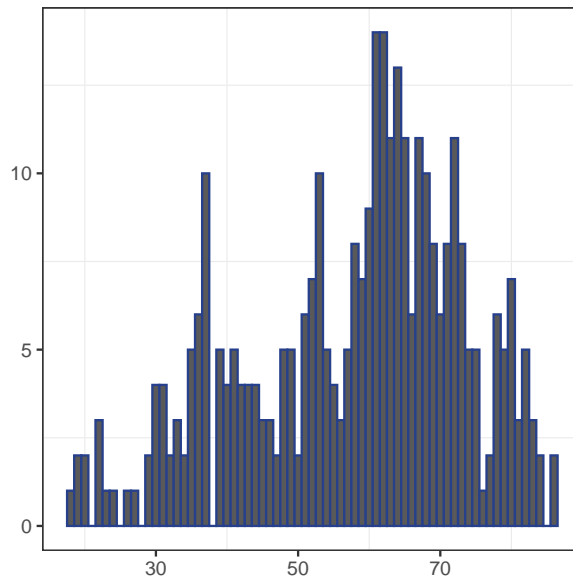
**Party one would vote for if elections were on Sunday**

Party	Counts	Percentage
KW Polskie Stronnictwo Ludowe (PSL, Kukiz15)	46	6.66
KW Prawo i Sprawiedliwo (PiS, Porozumienie, SP)	337	48.77
KW Sojusz Lewicy Demokratycznej (SLD, Wiosna, Lewica Razem)	77	11.14
KW Konfederacja Wolno i Niepodlego	31	4.49
KKW Koalicja Obywatelska PO .N IPL Zieloni	146	21.13
Other	12	1.74
Hard to say	26	3.76
Refuse to answer	16	2.32
Total	691	100

## Age Distribution



(a) Age distribution in the entire sample.



(b) Age distribution among PiS voters.

## Town Size

Town Size	Percentage	
	Entire Sample	PiS Voters
Village	39.72	46.09
Village - 20,000	16.55	14.82
20,000-50,000	11.31	8.08
50,000-100,000	9.24	5.92
100,000-500,000	15.45	11.05
500,000 and more	7.72	4.85

**Occupation**

Occupation	<i>Percentage</i>	
	Entire Sample	PiS Voters
Directors, presidents and managers of enterprises..	3.17	1.35
Creative professions and specialists with higher education, engineers,..	8.97	3.50
Technicians and other associate professionals, nurses, non-commissioned officers, policemen	3.72	2.96
Administrative and office workers, secretaries, postal workers, reception desk	5.66	4.31
Employees of shops, service points, employees of personal services	5.66	5.12
Skilled workers and foremen employed outside agriculture and forestry	10.48	11.05
Employees performing simple work outside agriculture and forestry	2.21	2.43
Salaried workers and foremen employed in agriculture and forestry	0.28	0.54
Individual farmers and family members assisting them	3.17	4.31
Owners and co-owners of private companies, plants, shops	4.14	1.08
Pupils, students	3.45	1.35
On disability insurance	4.97	7.28
Pensioners	36.00	38.54
Unemployed	1.24	1.08
Housekeepers, housewives	5.24	4.85
Not working for other reasons	1.66	1.08

## Education

Education	Percentage	
	Entire Sample	PiS Voters
Incomplete primary education or no school education	0.14	0.27
Basic	12.00	19.14
Junior high school	2.34	0.81
Basic Vocational (also SPR)	24.41	27.49
Secondary general education without Matura exam	3.03	2.96
Secondary general education with high school diploma	5.79	3.50
Secondary vocational without high school diploma	6.07	6.20
Secondary vocational with high school diploma	16.97	15.09
Post-secondary or post-secondary	2.90	1.35
Higher ed with the title of engineer, bachelor, certified economist	5.79	4.04
Higher degrees with a master's, doctor's degree or equivalent	19.45	9.70
Higher with a doctoral degree or higher	1.10	0.27

## EXPERIMENT RESULTS

We estimated the following model:

$$(11) \quad y_i = \alpha + \tau_1 \text{Treat}_i + \sum_g \gamma_g I_i\{\text{Age Group } g\} + \sum_g \tau_g \text{Treat}_i \cdot I_i\{\text{Age Group } g\} + \beta X_i + \epsilon_i$$

where  $\text{Treat}_i \in \{0, 1\}$  indicates whether respondent  $i$  was randomly assigned to the treatment or the control question.  $I_i\{\text{Age Group } g\} \in \{0, 1\}$  is an indicator function for whether respondent  $i$  belongs to age group  $g$ : age group 2 are 30-55 years-old and age group 3 are above 55 years of age, while our baseline group are 18-30 years-old. Standard errors are clustered at the Wojewdztwo level. Table A1 reports the results. In one specification we include a set of covariates  $X_i$ , which does not significantly change our results.

To check whether our results are a function of our choice of age groups, we re-estimate (11) with different choices of upper age threshold for the baseline group (the relevant one for our analysis). The estimate as well as 90 and 95% confidence intervals are reported in Figure A3.

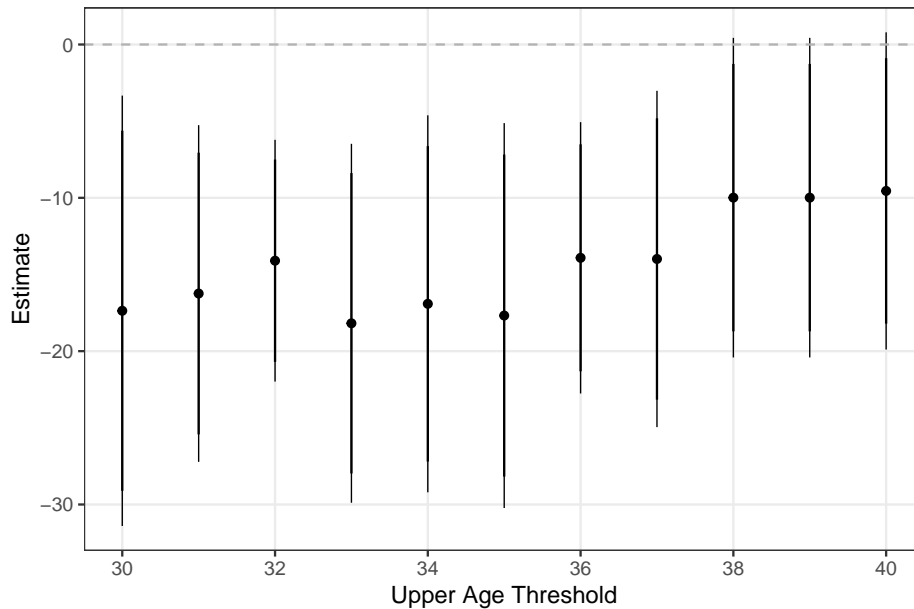
	<i>Dependent variable:</i>	
	Propability of voting for PiS	
	(1)	(2)
Treatment (Age Group 1)	-17.366** (7.160)	-17.005** (7.023)
Treatment x Age Group 2	15.232* (7.272)	14.821* (7.385)
Treatment x Age Group 3	16.863** (7.771)	15.993* (7.860)
Age Group 2	-7.211*** (1.937)	-6.697*** (2.076)
Age Group 3	-4.446 (2.837)	-3.746 (2.683)
Town Size		0.093 (0.538)
Higher Education		-0.286 (2.605)
Gender		-0.864 (1.608)
Religiosity		16.011 (16.835)
Unemployed		-0.944 (7.487)
Constant	97.689*** (1.658)	82.623*** (15.031)
Clusters	Wojewdztwo	Wojewdztwo
Observations	337	337
R <sup>2</sup>	0.036	0.052
Adjusted R <sup>2</sup>	0.021	0.023
Residual Std. Error	15.525 (df = 331)	15.51 (df = 326)
F Statistic	3.31** (df = 5; 331)	5.35*** (df = 10; 326)

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

**Table A1.** Treatment effect on the probability of voting for PiS, by age group. The three age groups are respectively: 18-30 year-olds, 30-55 year-olds, and those above 55.



### Robustness to Age Threshold



**Figure A3.** *Sensitivity of the treatment effect to changing the upper age threshold of the first age group (90 and 95% confidence intervals).*