Drain the Swamp: A Theory of Anti-Elite Populism

Gabriele Gratton and Barton E. Lee*

Latest version here

February 19, 2023

Abstract

We study a model of popular demand for anti-elite populist reforms that *drain the swamp*: replace experienced public servants with novices that will only acquire experience with time. Voters benefit from experienced public servants because they are more effective at delivering public goods and more competent at detecting emergency threats. However, public servants' policy preferences do not always align with those of voters. This tradeoff produces two key forces in our model: public servants' incompetence spurs disagreement between them and voters, and their effectiveness grants them more power to dictate policy. Both of these effects fuel mistrust between voters and public servants, sometimes inducing voters to drain the swamp in cycles of anti-elite populism. We study which factors can sustain a responsive democracy or induce a technocracy. When instead populism arises, we discuss which reforms may reduce the frequency of populist cycles.

*Gratton: UNSW Business School, UNSW Sydney. Email: g.gratton@unsw.edu.au. Lee: Dep. of Management, Technology, and Economics, ETH Zurich. Email: bartonlee@ethz.ch. We thank Benjamin Blumenthal, Allan Drazen, Ethan Kaplan, Federico Masera, and Massimo Morelli. Gratton is the recipient of an Australian Research Council Future Fellowship (FT210100176).

1 Introduction

Modern economies rely upon experienced state bureaucracies to effectively deliver public goods and services and competently adjust policy to a changing world. In democracies, voters choose representatives to work with the state bureaucracy and oversee its operations. But sometimes they elect populist leaders who *drain the swamp*: replace experienced public servants with less effective and incompetent novices.[1] For example, scholars in public administration and political science have noted how President Trump "sidelined administrative expertise and scientists in many areas, selecting senior leaders whose lack of qualification is frequently matched only by their disdain for their organizational mission" (Bauer, Peters, Jon, Yesilkagit and Becker, 2021).[2] Draining the swamp spurs widespread criticism from elite commentators who maintain that bureaucrats are humble servants of the voters, unfairly dismissed and replaced by less competent novices. However, outside elite circles, many voters applaud the dismissal of experienced public servants, arguing that elite public servants are too powerful and their policies disagree with what voters really need.[3]

Since draining the swamp is costly for the economy, voters' demand for it is puzzling. In some cases, voters may indeed become informed that public servants have reported misleading information and attempted to implement policies that voters did not want. For example, media coverage of revolving doors between financial regulator offices and banks in the wake of the Great Financial Crisis fueled widespread mistrust of governments and demand for change. But in other cases, voters demand the removal of elite public servants that they do not trust but that, according to expert commentators, are serving them well. For example, push-back against expert advice at the beginning of the Covid pandemic was widely condemned as unjustified mistrust of public health experts.

We develop a theoretical framework to conceptualize the forces that create mistrust

[1] Conservatives, such as Ronald Reagan, also use the expression "drain the swamp." They mean it as a call for reducing waste and inefficiencies in the bureaucracy. Populist leaders mean it as a call for a change of personnel: elite public servants should be replaced by personnel drawn from outside the bureaucracy.

[2] Moynihan (2021) discuss several high-profile cases. Beyond such cases, President Trump also appointed Ben Carson to the post of Secretary of Housing and Urban Development, a man without any previous bureaucratic experience and who stated that having him "as a federal bureaucrat would be like a fish out of water" (Costa, 2016).

[3] As reported by Massimo Calabresi in Time magazine, Susan Rose Ackerman's (Yale Law School) view on Trump's war on the bureaucracy is that it is a "war on the core responsibility of the bureaucracy to make sure the laws as passed are carried out" Yet, "For many of Trump's followers [it is] exactly what they asked for [. . .] "He was given a mandate with the election to go up there and correct and fix Washington and drain that swamp. That is exactly what I see him doing," says Janice Westmoreland, 69, of Milledgeville, Ga" (Calabresi, 2017). More favorable reviews of President Trump's action notice that "for years, unelected bureaucrats have been allowed largely unchecked power over the daily lives of Americans. This president is trying to change that" (Bovard, 2019).

between voters and public servants. We argue that voters need experienced public servants to effectively provide public goods and competently detect and respond to new threats and opportunities. However, when threats are uncertain, public servants' preferred response may not perfectly align with voters' preferences. For example, voters and public servants alike may prefer a policy of fiscal austerity when financial instability is sufficiently likely. But elite public servants pay a relatively smaller cost for such policies as their jobs are more secure. As a result, compared to voters, public servants may prefer to enact austerity policies at lower levels of financial instability risk.

In Section 2, we formalize this idea in a dynamic agency model. In each period, the voter inherits an incumbent agent who is either *experienced* or a *novice* that will acquire experience over time. An experienced agent is both more *effective* in providing public goods[4] and *competent*: she can detect threats that should be countered by triggering costly emergency policies. From the point of view of the voter, the optimal policy—what we refer to as the agent's *mandate*—is to trigger emergency policies if and only if the agent detects a threat. However, agents are more hawkish than voters: their cost of triggering an emergency policy is smaller. Whether she detects a threat is an experienced agent's private information, but if she triggers an emergency policy, oversight institutions that enforce *transparency* (e.g., the media or parliamentary inquiries) may reveal the agent's information to the voter. The only tool available to the voters to avoid a triggered emergency policy is to replace the agent—to drain the swamp. Draining the swamp may serve to discipline the agent but is costly for the voter: it replaces an experienced agent with a novice that will only acquire effectiveness and competence over time.

While our focus is on the rise of anti-elite populism, we believe this model may provide useful insights in other contexts. For example, our model finds a natural application in the study of agency problems between a firm's board and its manager. Like public servants, managers also acquire firm-specific experience with time. Experienced managers are both more effective in managing the normal operations of the firm and more competent in detecting opportunities in new markets that fit the firm's potential and design products that seize such opportunities. Both the board and the manager wish to seize opportunities when they arise but know that attempting to seize false opportunities is costly. However, the manager's career is further boosted if she successfully seizes an opportunity, so that, compared to the board, she is more averse to missing a true opportunity than seizing a false one. When the manager pursues an aggressive policy in new markets, the

[4]As we discuss later, effectiveness may also include the agent's performance in any other task—including responding to other emergencies—in which disagreement between voters and public services does not arise.

board can choose to fire her (draining the swamp—in our language), but this is costly, as it replaces an experienced manager with a novice who is both less effective and less competent and will only acquire experience over time.

We show in Section 3 that our theoretical framework produces two key forces that determine equilibrium behavior. First, we show that *incompetence spurs disagreement*: only sufficiently incompetent agents ever have an incentive to deviate from their mandate and trigger emergency policies without having detected a threat. In fact, a very competent agent would detect threats with very high probability, and thus has no reasons to trigger an emergency policy when she does not detect one. Second, we show that *effectiveness begets power* to the agent: more effective agents are more costly to replace and thus need not to bother with the voter's threat to drain the swamp.

These two effects drive the conditions that determine which regime arises in the unique equilibrium we characterize in Section 4. A sufficiently powerful agent establishes a *technocracy*, whereby she dictates policymaking at will, independently of the voter's mandate. Otherwise, less intense disagreement between the voter and the agent supports a *responsive democracy* in which the agent always abides by her mandate. But more intense disagreement induces cycles of anti-elite populism, whereby the agent violates her mandate and the voter drains the swamp along the equilibrium path.

We show that the power of the agent determines which of two types of populist cycles arise. In *informed populism*, the voter only drains the swamp when he becomes informed that the agent has indeed violated her mandate and acted against the voter's interest. In contrast, in *preemptive populism*, the voter drains the swamp with positive probability even when uninformed. In fact, an expert commentator, able to access and assess the information available to the agent, would sometimes conclude that the voter is replacing an agent he does not trust but that is serving him well.

Our theory allows us to shed some light on both the potential reasons for why populism is on the rise in the past decades and what, if anything, can be done to mitigate its frequency and impact. In Section 5 we argue that our model suggests that well-functioning democracies may plunge into populism as a result of four possible shocks. A decrease in the competence of public servants, an increase in their hawkishness or in the frequency of threats all can induce populist cycles by exacerbating disagreement. In addition, populist cycles may also be brought upon by a decrease in the transparency of institutions that breaks the trust relation between voters and public servants. We relate these four shocks to social, political, and technological trends in Western democracies in the last four decades.

Our model sharpens the intuition that managing populism is a difficult task. We show in Section 6 that the optimal response depends on several factors, including whether the size of the shock that induced populism in the first place is reversible. When this is not possible, a reformer may attempt to at least reduce the frequency of populist cycles in which voters drain the swamp. In this case the optimal policy depends on whether the country has plunged into an informed or preemptive populist regime. Reducing the frequency of cycles of informed populism—perhaps counterintuitively—requires reforms that decrease transparency and select less competent public servants. Reducing cycles of preemptive populism, instead, requires increasing transparency or selecting less hawkish public servants. Changing public servants' competence may instead backfire as it may both increase or decrease the frequency with which voters drain the swamp. In fact, the only policy that unambiguously (albeit weakly) reduces the frequency of populist cycles is a reduction in the underlying cause of the public servants' hawkishness—for example, by reducing labor protection differences between private and public sectors, or selecting bureaucrats with preferences that are more aligned with those of the average voter. Thus, our theory suggests that the most robust remedy against anti-elite populism is the design of a more inclusive and representative bureaucratic elite, perhaps through the design of meritocratic systems that favor the selection of personnel from a broader and more inclusive set of socio-economic backgrounds. Finally, our model warns against the technocratic peril of attempting to solve anti-elite populism by increasing the effectiveness of public service.

In Section 7 we discuss how our results change (or do not change) when we vary our assumptions about what the agent cares about and how long it takes her to accumulate experience. We believe these extensions to be useful in other applications beyond our main focus on anti-elite populism. Section 8 concludes the paper.

1.1 Related literature

Our theory focuses on voters' demand for anti-technocratic and anti-bureaucratic reforms. Populism is an elusive concept to define and not all populist movements drain the swamp.[5] However, mistrust for the elites governing democratic policy-making is common to many populist movements. In fact, Mudde and Rovira Kaltwasser (2017) define populism as an ideology that separates society into "the pure people" and the "the corrupt elite" and, in contrast to elitism, expresses the view that the will of the pure peo-

[5] For example, some populist leaders, such as Ecuadorian president Rafael Correa, have taken technocratic approaches to governing—promising to utilize 'outsider' expertise to make the bureaucracy more competent and more effective (see, e.g., Bauer et al., 2021; Panizza, Peters and Ramos Larraburu, 2019; Postel, 2007).

ple should trump that of the elite. We abstract from the source of a supply for populist leaders and focus on what drives voters to demand that elite bureaucrats are replaced with less experienced ones who implement the people's will. Our model provides insights into when disagreement is more likely to arise between the people and the elites, and when the elites may become powerful enough to dictate policy. We argue that a well-functioning democracy governs disagreement so that the elite serves the people well. In contrast, when the elite is less competent, populism arises as an ideology that prefers inexperienced public servants to elite bureaucrats.

In our model, inexperienced public servants implement default policies while experienced ones claim to competently enact policies that respond to a changing world. We share this view of populism with Morelli, Nicolò and Roberti (2021), who argue that demand for populism arises from a desire for politicians who commit to simpler policies that can be more easily monitored.

Demand for populism may come for other reasons: Economic insecurity or other threats to voters' welfare (Algan, Guriev, Papaioannou and Passari, 2017; Ananyev and Guriev, 2019; Autor, Dorn, Hanson and Majlesi, 2020; Colantone and Stanig, 2018; Gratton and Lee, forthcoming; Guiso, Herrera, Morelli and Sonno, 2017, 2019, 2020; Rodrik, 2018), or other sources of mistrust in institutions (Norris and Inglehart, 2019) (see Guriev and Papaioannou, 2020, for a review). Our analysis complements this literature by providing a mechanism through which populism addresses voters' economic and cultural concerns. In particular, in our model, greater economic insecurity increases the disagreement between the bureaucratic elite and the voters, thus increasing the demand for leaders who promise to drain the swamp. Furthermore, our theory endogenizes the source of mistrust between voters and bureaucrats, identifying under which circumstances elite bureaucrats disagree with voters and the determinant of their power to dictate policy.

An emerging literature has documented the negative effects of populist governments on the quality of national bureaucracies (Bellodi, Morelli and Vannoni, forthcoming; Moynihan, 2021). Sasso and Morelli (2021) study a model in which populist politicians implement reforms that decrease the quality of the bureaucracy because they prefer bureaucrats who implement their platforms. Our theory provides insights into why voters may demand such costly reforms and elect leaders who promise to drain the swamp and implement the policy voters want, independently of the expert advice of experienced public servants.

Our paper contributes to a broader literature on bureaucratic control in democracies (see Gailmard and Patty, 2012a, for an overview) and how politics affects the quality of the bureaucracy (e.g., Gratton, Guiso, Michelacci and Morelli, 2021; Nath et al., 2016;

Ting, 2021). Implicit in our theory is the idea that bureaucrats have discretionary power that voters and politicians can control only with blunt instruments such as replacing personnel. We share this idea with, e.g., Banks and Weingast (1992) and Bendor and Meirowitz (2004). Gailmard and Patty (2012b) highlight that bureaucrats are given discretionary power because it creates incentives to accumulate expertise (see also Alesina and Tabellini, 2007, 2008; Callander and Krehbiel, 2014; Huber and Shipan, 2002; Maskin and Tirole, 2004; Ting, 2002, who provide other explanations for the discretionary power of bureaucrats). We depart from this literature in two directions. First, we explicitly model the observation that "expertise development takes time" (Gailmard and Patty, 2012b, p. 26). Second, we highlight the interaction between two features of expertise—competence and effectiveness—and show that each plays a distinct role in the relationship between voters and public servants.

Our model shares some features with models of both bureaucratic control and political accountability stemming from the seminal work of Holmström (1980, 1982), Barro (1973), and Ferejohn (1986) (e.g., Huber and McCarty, 2004; Morris, 2001; Dewan and Squintani, 2018; Fox and Jordan, 2011; Besley, 2006; Fox and Shotts, 2009; Ashworth, Bueno de Mesquita and Friedenberg, 2018; Ashworth and Bueno de Mesquita, 2014; Kartik, Van Weelden and Wolton, 2017; Duggan and Martinelli, 2017). Most of this literature focuses on disagreement that arises between a principal and an agent because the principal and the agent want different things. In contrast, we assume that voters and public servants essentially want the same thing, but when there is uncertainty about which option is best, disagreement arises because of their differing tolerances for different types of mistakes. This view yields our result that incompetence spurs disagreement.

Our formal model abstracts from the role of political leaders. In this sense, we offer a theory of the relation between the citizens and the state. It posits that robust, well-functioning democratic institutions only survive in a "narrow corridor" in which the state is sufficiently competent but not too powerful (echoing arguments in Acemoglu and Robinson, 2019; Stasavage, 2020) and that democracy is inherently vulnerable to technological and social shocks that increase disagreement between different classes of citizens (Przeworski, 2019).

2 The model

2.1 Summary

We study a model with a forward-looking and infinitely lived voter. In each period, the voter either inherits an effective and competent state organization with experienced public servants or one with inexperienced novices. We call the organization of the state an *agent* and say that the agent is either *experienced* or a *novice*. Compared to a novice, an experienced agent is more *effective* at producing public goods and services, as well as more *competent* at detecting emergency threats.

We model the experienced agent's competence as the precision of a signal that she observes as the new period begins. The signal is binary: she either detects a threat or not. The agent can then trigger an *emergency policy*. The voter's preference is for the agent to trigger the emergency policy if and only if she has detected a threat. We refer to this strategy as the *agent's mandate*. However, agents are *hawkish*: compared to the voter, agents have a greater aversion to emergencies that are not covered by emergency policies so that, for some parameters of the model, an agent may prefer to trigger the emergency policy even when she detects no threat. When this is the case, we say that the voter and the agent *disagree*.

Triggering an emergency policy raises public attention to possible threats so that the voter has a chance to observe directly the information available to the agent. This chance naturally increases with the quality and *transparency* of the bureaucracy, the media, and the public debate.

Whether the voter observes the information or not, she can avoid the emergency policy by *draining the swamp*: replace the incumbent, experienced agent with a novice, unable to detect threats or implement emergency policies. Draining the swamp is costly as it deteriorates the ability of the state to produce goods and services and, for the time being, generates a state unable to detect threats and respond to them with appropriate emergency policies. However, novices accumulate experience over time so that, after a number of periods, a novice becomes experienced.

2.2 Formal setup

A forward-looking voter lives for infinitely many periods $t \in \{1, 2, \dots\}$. In each period t, there is an incumbent agent who is either *experienced* or a *novice*, and either an *emergency* occurs, $\theta_t = 1$, or not, $\theta_t = 0$, with $\pi \equiv \Pr(\theta_t = 1) < 1/2$.[6]

[6]The assumption that $\pi < 1/2$ means that emergencies are less likely than non-emergencies.

An experienced agent privately observes a binary signal $s_t \in \{0, 1\}$ about whether an emergency has occurred, where $s_t = 1$ means that the agent *detects an emergency threat*. The *competence* of an experienced agent is captured by the precision of the signal s_t, $\kappa \equiv \Pr(s_t = \theta_t \mid \theta_t) > 1/2$.

Upon observing the signal, the agent chooses whether to trigger an *emergency policy*, $p_t = 1$, or not, $p_t = 0$. If the agent triggers an emergency policy, then the voter observes the agent's signal s_t with probability equal to the *transparency* of the system $\tau > 0$. Whether she observes the signal or not, the voter chooses whether to *drain the swamp*: replace the experienced agent with a *novice* that does not detect threats and never triggers emergency policies.[7] Let $d_t = 1$ denote the decision to drain the swamp in period t and $d_t = 0$ its complement. The policy implemented equals $i_t = p_t(1 - d_t)$.

Agents accumulate experience over time so that an agent first hired at time t is a *novice* at time $s \geq t$ if $s < t + T$ and *experienced* thereafter. In our benchmark model we assume $T = 1$ so that a novice becomes experienced in one period. For simplicity, we assume that at $t = 1$ the incumbent agent is experienced.

The voter discounts future payoffs with factor $\delta < 1$. Her period-t payoff is given by the sum of the public goods produced in that period and a payoff equal to 1 if $i_t = \theta_t$ and 0 if $i_t \neq \theta_t$. The extra amount of public goods and services afforded by an experienced agent compared to a novice equals the experienced agent's *effectiveness*, $\eta > 0$.

If the voter does not drain the swamp in period t, the incumbent agent receives a policy payoff and a continuation payoff; otherwise her payoff is 0. In our benchmark model, the agent is myopic so that the continuation payoff equals $\delta V > 0$.[8] Like the voter, the agent also prefers to implement emergency policies only in an emergency. However, the agent is *hawkish*: relative to the voter, the agent has a greater opportunity cost, normalized to 1, when an emergency policy is not triggered in an emergency (a type-II policy error) than her opportunity cost $1 - \alpha < 1$ when an emergency policy is triggered without an emergency (a type-I error). Naturally, a greater α means that the agent is more hawkish. Summarizing, if the voter does not drain the swamp in period t, the incumbent agent receives a period-t payoff equal to $1 + \delta V$ if $i_t = \theta_t = 1$, $(1 - \alpha) + \delta V$ if $i_t = \theta_t = 0$, and δV if $i_t \neq \theta_t$. Otherwise her payoff equals 0.

To focus on interesting cases, we assume that an experienced agent is sufficiently com-

[7] As shall be clear, the assumption that a novice agent is a non-strategic player can be naturally microfounded. For example, if a novice agent has no competence ($\kappa = 1/2$) and no effectiveness ($\eta = 0$), then in equilibrium the voter would drain the swamp whenever a novice triggers the emergency policy and a novice agent would never trigger it.

[8] In Section 7 we discuss an extension of the model in which the agent is forward-looking so that V is endogenously determined by the expected present discounted sum of policy payoffs for the agent until the voter drains the swamp and dismisses her (see also, Appendix D).

petent so that the *agent's mandate*, if carried out optimally for the voter, is to choose a policy equal to the signal observed by the agent: $\kappa > 1 - \pi$.

Assumption 1 (The agent's mandate.)**.** *It is optimal for the voter to implement a policy equal to the signal observed by an experienced agent:* $\kappa > 1 - \pi$.

We characterize the perfect Bayesian equilibrium in Markovian strategies that survive divinity (henceforth, equilibrium).[9] In Appendix A, we provide a formal definition of the equilibrium. With the exception of knife-edge cases, this equilibrium is unique. All proofs are in Appendix B.

3 Disagreement and power

In this section, we study the key tradeoffs faced, in each period, by the strategic players of our model: the voter and the incumbent experienced agent. We show how the parameters of the model affect two key endogenous tensions between the voter and the agent: how much the agent and the voter *disagree* about policymaking and the agent's *power* to dictate policy. These two tensions will drive the characterization of the equilibrium in Section 4.

3.1 Competence and disagreement

In our model, the voter and the agent both want to match the policy with the state, but the informed party—an experienced agent—is more hawkish. However, this does not imply that, for a given realization of the signal s_t, the agent and the voter would disagree on the optimal policy. For example, were $\kappa = 1$ so that the agent is perfectly informed about the state, the voter and the agent would never disagree, for any level of hawkishness α. Lemma 1 says that disagreement arises if and only if the agent is sufficiently incompetent.

Lemma 1 (Incompetence spurs disagreement.)**.** *There exists* $\bar{\kappa}(\alpha, \pi)$ *such that if* $\kappa > \bar{\kappa}(\alpha, \pi)$*, then both the voter and an experienced agent strictly prefer to implement the emergency policy if and only if the agent detects a threat; otherwise the two disagree and the agent prefers to implement the emergency policy independently of whether she detects a threat.*

Intuitively, a sufficiently incompetent agent is less confident in her signal. Therefore, even upon observing $s_t = 0$, she is still primarily concerned by the possibility of a type-II policy

[9] Divinity (Banks and Sobel, 1987; Cho and Kreps, 1987) is a standard refinement in the signaling literature. In our context it requires the voter to attribute a deviation to triggering an emergency policy (when in equilibrium the agent is always expected not to do so) to the type of informed agent who would choose it for the widest range of voter's responses.

error and prefers to trigger the emergency policy. However, were the voter to observe the same signal, he would instead prefer not to implement the emergency policy. Naturally, the disagreement threshold $\bar{\kappa}(\alpha, \pi)$ depends on both how hawkish the agent is and the underlying frequency of emergencies. So while incompetence spurs disagreement, its effects are exacerbated by other elements in our model that contribute to disagreement: the agent's hawkishness and the frequency of emergencies.

From now onward, we focus on the interesting case in which there is disagreement.

Assumption 2 (Disagreement.)**.** *The voter and an experienced agent disagree:* $\kappa < \bar{\kappa}(\alpha, \pi)$.

As we show in Appendix C, whenever this assumption is violated, in the unique equilibrium the agent optimally carries out her mandate and the agent never drains the swamp along the equilibrium path. Instead, when there is disagreement, it is possible that the agent may prefer to *violate her mandate*: choose to trigger the emergency policy even when she does not detect a threat (when $s_t = 0$).

In reality, public servants carry out multiple tasks, triggering a variety of emergency policies in response to differing threats. Disagreement is likely to arise only in a subset of these tasks. Our model may be interpreted as follows. Effectiveness η captures the payoff of having an experienced public service optimally carrying out their mandate on all tasks on which disagreement never arises. Assumption 2 can then be interpreted as assuming that there exists one task in which, for some information available to the public service, voters and public servants disagree.

3.2 Drain the swamp

We now describe the tradeoff faced by the voter when the agent triggers the emergency policy. The voter needs to choose whether to let the agent implement the emergency policy or to drain the swamp. Draining the swamp is costly for two reasons. First, it replaces an experienced agent with a less effective novice that produces η fewer public goods. Second, the agent may have triggered the emergency policy because she has competently detected a threat ($s_t = 1$). In this case, the voter would indeed prefer to implement the emergency policy. However, draining the swamp may also be beneficial. The agent may have triggered the emergency policy but there is no actual emergency ($\theta_t = 0$).

Let ν_t be the voter's belief that an emergency has occurred. Lemma 2 says that the voter optimally drains the swamp whenever ν_t is sufficiently small and the agent is not too effective.

Lemma 2 (Voter's optimal strategy.)**.** *In any equilibrium, the voter drains the swamp if and only if*

$$\nu_t < \frac{1-\eta}{2}.$$

The voter's belief ν_t that an emergency has occurred depends both on the agent's strategy and on whether the voter observes the agent's signal s_t. Naturally, if the voter observes the agent's signal, her belief is greatest when the signal equals 1 and smallest when it equals 0. Lemma 3 says how the voter optimally chooses in these two extreme cases.

Lemma 3 (The optimal choice of an informed voter.)**.** *In any equilibrium, if the voter observes the agent's signal s_t, then*

(i) *when $s_t = 1$, the voter never drains the swamp;*

(ii) *when $s_t = 0$, there exists $\bar{\eta}(\pi, \kappa)$ such that the voter drains the swamp if and only if $\eta < \bar{\eta}(\pi, \kappa)$.*

Intuitively, if the voter is informed that the agent has indeed detected a threat ($s_t = 1$), then the voter knows that the agent has optimally carried out her mandate and he prefers not to drain the swamp. If instead the voter is informed that the agent has not detected a threat ($s_t = 0$), then the voter knows that the agent has violated her mandate. He then chooses to drain the swamp if and only if the agent is sufficiently ineffective, so that the cost of replacing her with a novice is not too large. Notice that this means that a sufficiently effective agent can expect not to be replaced even when the voter is informed that the she has indeed violated her mandate. As we shall see, this affords an effective agent actual *power* to dictate policy.

When the voter is not informed of the agent's signal, her choice depends on her trust in the agent. If the voter trusts the agent to abide by her mandate, then the voter should never drain the swamp: if the agent has triggered the emergency policy, then it is because she has detected a threat and responded optimally. However, because the agent may disagree with the voter, the voter may not trust her and believe that she has triggered the emergency policy without having detected a threat. This lack of trust in the agent is what may trigger the voter to *preemptively* drain the swamp even if she is *not* informed.

3.3 Effectiveness and power

Whether the agent can be trusted to abide by her mandate depends on her choice when she does not detect a threat ($s_t = 0$). By Assumption 2, the agent prefers to implement the emergency policy. Therefore, she would like to trigger the emergency policy so that she may be able to implement it if the voter does not drain the swamp. But triggering the emergency policy is risky: the voter may drain the swamp so that the agent is replaced.

Lemma 4 says that this tradeoff is resolved in two steps. First, a sufficiently effective agent does not bother with the voter because the voter will not drain the swamp even if he is informed that the agent violated her mandate. We say that in this case the agent is "powerful" enough to dictate policy and violate her mandate if she wishes to do so. Otherwise, the agent trades off the costs and benefits of triggering an emergency policy, including the possibility that the voter will drain the swamp. The voter may do so either because, with probability τ, he becomes informed that the agent has violated her mandate or because he does not trust the agent and chooses to preemptively drain the swamp.

Lemma 4 (Effectiveness begets power.)**.** *Let $\mu(s_t)$ be the agent's belief that an emergency has occurred when she observes signal s_t. In any equilibrium, if $\eta > \bar{\eta}(\pi, \kappa)$, the agent always triggers the emergency policy. If $\eta < \bar{\eta}(\pi, \kappa)$, the agent triggers the emergency policy with certainty when she detects a threat and otherwise with strictly positive probability only if*

$$(1 - \alpha)(1 - \mu(0)) + \delta V \leq (1 - \tau)(1 - d^*)(\mu(0) + \delta V)$$

where d^* is the equilibrium probability that the voter preemptively drains the swamp when she is not informed.

The last inequality in Lemma 4 is intuitive. When the agent does not detect a threat, she believes that an emergency has occurred with probability $\mu(0)$. Therefore, her expected policy payoff from implementing the emergency policy is $\mu(0)$ while her expected payoff from not implementing it is $(1 - \alpha)(1 - \mu(0)) < \mu(0)$. However, because the agent is not powerful enough, the voter may drain the swamp if the agent triggers the emergency policy. So, not triggering the emergency policy yields to the agent a less preferred policy but guarantees the continuation payoff V. Triggering the emergency policy results in a more preferred policy and the continuation payoff if the voter does not drain the swamp—otherwise the agent's payoff is 0. The probability with which the voter drains the swamp depends on the transparency of the system τ, which leads the voter to be informed, and the probability d^* with which the voter preemptively drains the swamp.

4 Technocracy, democracy, and populism

We now discuss how the two forces that we have identified—that incompetence spurs disagreement and that effectiveness begets power—drive equilibrium behavior and induce one of four different regimes.

Technocracy. In this regime, in every period, the agent violates her mandate: she always triggers the emergency policy. The agent is so powerful that she does not need to bother with the voter: the voter never drains the swamp, even if he becomes informed that the agent violated her mandate.

Responsive democracy. In this regime, in every period, the agent abides by her mandate and the voter drains the swamp if and only if he becomes informed that the agent violated her mandate. On the equilibrium path, the voter never drains the swamp.

Informed populism. In this regime, in every period, the agent violates her mandate: she always triggers the emergency policy. The voter drains the swamp if and only if he becomes informed that the agent violated her mandate. On the equilibrium path, the voter periodically drains the swamp with probability $\tau \Pr(s_t = 0)$.

Preemptive populism. In this regime, in every period, the agent violates her mandate: she triggers the emergency policy with certainty when she detects a threat and with probability $p^* > 0$ otherwise. The voter drains the swamp both when informed and preemptively: he does so with certainty when he becomes informed that the agent violated her mandate and with probability $d^* > 0$ otherwise. On the equilibrium path, the voter periodically drains the swamp with probability $\tau \Pr[s_t = 0]p^* + (1 - \tau)d^*$.

Proposition 1 characterizes the unique equilibrium. It says that a sufficiently powerful agent establishes a technocracy. Otherwise, less intense disagreement between the voter and the agent supports a responsive democracy. But when disagreement is more intense, the voter drains the swamp on the equilibrium path and the power of the agent determines whether cycles of informed or preemptive populism arise. Figure 1 provides an illustration.[10]

Proposition 1 (Technocracy, democracy, and populism.)**.** *There exists cutoffs $\underline{\kappa}(\pi, \alpha, \tau) < \bar{\kappa}(\pi, \alpha)$ and $\underline{\eta}(\pi) < \bar{\eta}(\pi, \kappa)$ such that, in the unique equilibrium*

(i) *$\bar{\eta}(\pi, \kappa) < \eta$ induces a technocracy;*

(ii) *$\eta < \bar{\eta}(\pi, \kappa)$ and $\underline{\kappa}(\pi, \alpha, \tau) < \kappa$ induces a responsive democracy;*

(iii) *$\underline{\eta}(\pi) < \eta < \bar{\eta}(\pi, \kappa)$ and $\kappa < \underline{\kappa}(\pi, \alpha, \tau)$ induces informed populism;*

[10] For completeness, we recall that when Assumption 2 is relaxed, so the voter and the agent do not disagree, the agent carries out her mandate optimally and, on the equilibrium path, the voter never drains the swamp. This regime is observationally equivalent to a "responsive democracy."

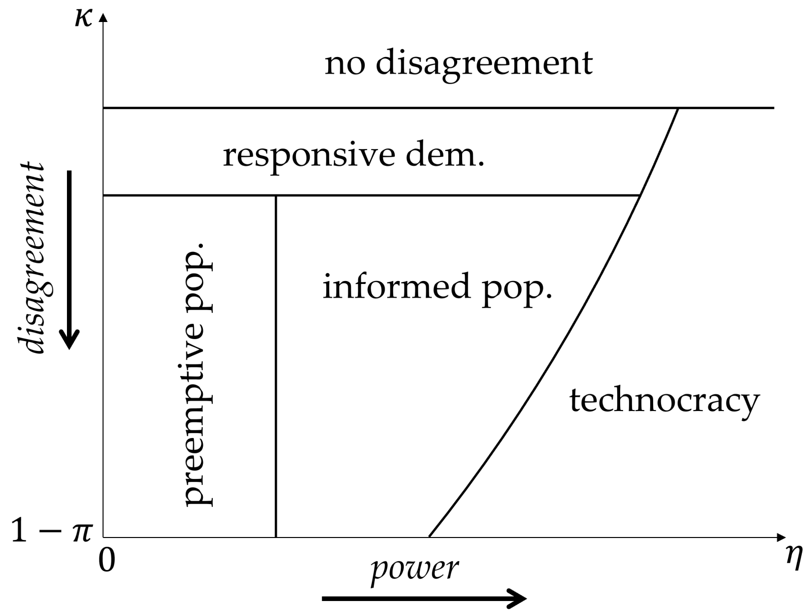

Figure 1: Typology of regimes. Parameter values: $\pi = 0.4, \tau = 0.35, \delta = 0.05, \alpha = 0.95, V = 1$.

(iv) *$\eta < \underline{\eta}(\pi)$ and $\kappa < \underline{\kappa}(\pi, \alpha, \tau)$ induces preemptive populism.*

Intuitively, when the agent is very effective, she is sufficiently powerful that she can dictate policy without fear of being replaced, giving rise to a technocracy. Otherwise, the agent's competence determines whether a responsive democracy may be sustained. In a responsive democracy, the voter drains the swamp when he becomes informed that the agent violated her mandate. This induces the agent to abide by her mandate, but only if she does not disagree too intensely. Therefore, only a sufficiently competent agent can sustain a responsive democracy. When the agent is less competent, so that disagreement is more intense, her effectiveness matters in determining behavior. This is because when the agent is more effective, the voter is only willing to drain the swamp when she becomes informed that the agent violated her mandate. This empowers the agent to always trigger the emergency policy. When the agent is less effective, the voter is in addition willing to preemptively drain the swamp so to discipline the agent and the agent only sometimes violates her mandate.

A key feature that characterizes preemptive populism is that the voter decision to drain the swamp may result in policies that are worse for the voter. In fact, when the voter drains the swamp preemptively, it may well be the case that the agent has not violated her mandate: she triggered the emergency policy because she has indeed detected a threat. A competent external observer who can see the agent's signal will therefore conclude that the voter is draining the swamp against his own interest.

A key lesson from Proposition 1 is that more intense disagreement causes populism. The agent's incompetence spurs disagreement, but its effect is exacerbated by the agent's hawkishness, α, and the frequency of emergencies, π. The combined effect of these three factors in determining disagreement underscores the results in the following two sections.

In these sections we show that our equilibrium characterization is a useful framework for addressing two sets of questions. First, we show that it provides insights into which shocks to the political and social environment may break the trust relation between voters and public servants. When this occurs, a well-functioning democracy—one in which public servants work in the interest of voters—may become dysfunctional: public servants routinely betray voters and voters do not trust public servants, plunging the country into cycles of anti-elite populism. Second, we show that our framework also helps in understanding which reforms, by manipulating the power of public servants and the level of disagreement between them and the voters, may alleviate the frequency of populist cycles when the trust relation between voters and public servants has broken.

5 The four horsemen of populism

Functional democracies sometimes develop strong popular demand for anti-elite populism. In this section we use our model to pin down which shocks to the political and social environment are likely to cause this shift. A key feature of our model is that even a shock that barely moves a country from a responsive democracy into a populist regime causes a discrete jump in the probability that the agent violates her mandate, thus breaking the trust relation between the voter and the agent.[11]

Proposition 2 identifies two types of shocks that, if sufficiently large, may plunge a responsive democracy into either an informed or a preemptive populist regime. One type of shocks induces populism by intensifying disagreement between the voter and the agent. This may be brought upon by either a fall in the agents' competence, an increase in their hawkishness, or more frequent emergencies. A second type of shock that may cause populism occurs when the transparency of the system drops.

Proposition 2 (The four horsemen of populism.)**.** *A responsive democracy may lead to informed or preemptive populism in response to shocks that decrease the agent's competence, κ,*

[11] In particular, in all populist regimes, at the limit as $(\kappa, \pi, \alpha, \tau)$ approach $\kappa = \underline{\kappa}(\pi, \alpha, \tau)$, the probability p^* that the agent triggers the emergency policy when not detecting a threat is bounded away from zero. In contrast, the probability d^* with which the voter preemptively drains the swamp approaches zero as the parameters approach this threshold.

increase her hawkishness, α, or the frequency of emergency threats, π, or decrease transparency, τ. In particular, if $\kappa, \alpha, \pi, \tau$ induce a responsive democracy, then:

(i) *There exists $\kappa' \in (1-\pi, \kappa)$ that induces populism if and only if $1-\pi < \underline{\kappa}(\pi, \alpha, \tau)$ and $\eta < \bar{\eta}(\pi, \underline{\kappa}(\pi, \alpha, \tau))$.*

(ii) *There exists $\alpha' \in (\alpha, 1)$ that induces populism if and only if $\kappa < \underline{\kappa}(\pi, 1, \tau)$.*

(iii) *There exists $\pi' \in (\pi, 1/2)$ that induces populism if and only if $\kappa < \underline{\kappa}(\pi_\eta, \alpha, \tau)$, where π_η is defined as the unique value $x \in (\pi, 1/2)$ such that $\eta = \bar{\eta}(x, \kappa)$ whenever it exists and, otherwise, $\pi_\eta = 1/2$.*

(iv) *There exists $\tau' \in (0, \tau)$ that induces populism if and only if $\kappa < \underline{\kappa}(\pi, \alpha, 0)$.*

Intuitively, a responsive democracy requires two essential conditions. First, the agent and the voter must not disagree too intensely, so that the agent's incentive to violate her mandate is not too strong. Second, the system must be sufficiently transparent, so that the agent is deterred from violating her mandate by the threat that the voter will become informed.

Naturally, each of the effects caused by the four shocks in Proposition 1 interact with each other. Points (ii)–(iv) say that if the agent is sufficiently competent, then even large shocks in the other three parameters cannot spur sufficient disagreement to cause populism. However, if the agents are not too competent, then there always exist a shock in α, π, or τ that will plunge a responsive democracy into a populist regime. The two further conditions in points (i) and (iii) guarantee that the shocks in κ and π lead to populism rather than technocracy (see Section 6).

Which type of populism is induced by each of these four shocks depends on the power of the agent. When the agent is more powerful, the voter's cost of draining the swamp preemptively is too large and the shocks in Proposition 2 result in informed populism. Otherwise, the agent drains the swamp preemptively.

Corollary 1 (Power and populism.). *Let $\kappa, \alpha, \pi, \tau$ induce a responsive democracy. If a shock in any of the four parameters leads to a populist regime, then the resulting regime is informed populism if $\eta > \underline{\eta}(\pi')$ and preemptive populism otherwise, where π' equals the post-shock value of π (including, possibly $\pi' = \pi$).*

Proposition 2 offers a theoretical framework to explain why anti-elite populism has risen in Western democracies in the first decades of the 21st Century. Globalization and the Great Financial Crisis have eroded the (at least perceived) ability of Western national bureaucracies to address the demand for economic and financial stability of the middle class, as well as increased the frequency of threats to economic stability. In our model, this corresponds to a drop in the agent's competence κ and an increase in π, increasing disagreement and reducing the voter's trust in the agent. Furthermore, labor reforms across the Atlantic starting in the last two decades of the 20th Century have increased the disparity in employment conditions between private and public sector, with unionization and permanent jobs increasingly present only in the public sector. This pattern has exacerbated disagreement between public servants—who are less likely to lose their job in a recession—and voters. In our model this corresponds to a more hawkish agent (higher α). Finally, the information technology revolution has exposed voters to a multitude of sources offering a cacophony of fake news, reducing the ability of voters to assess the effective risk associated with differing policies, and therefore reducing the accountability of public servants. In our model, this is captured by a drop in transparency τ. According to our model, each of these factors, in isolation or together, has the potential to plunge a responsive democracy into informed or even preemptive populist cycles.

6 Managing populism

In this section we discuss which reforms may be put in place to address the problem of anti-elite populism after the country has plunged into a populist regime. Naturally, *large* reforms that reverse the course of the shock (or shocks) that have led the country into populism will suffice. For example, by Proposition 2, if the country has plunged into populism because public servants have become less competent, then reforms that restore their previous level of competence (e.g., more effective selection of bureaucrats or strengthening collaborations with scientists who can help detect threats) will inevitably restore a responsive democracy. However, such large reforms may not be possible. For example, the fall in competence may be due to technological shocks or changes in the international environment that are beyond the reach of domestic reforms. In such cases, a reformer concerned about populism may be constrained to only implement *small* reforms that cannot eliminate populism altogether but may nonetheless at least reduce its frequency.

The following propositions inform the choice of a reformer concerned about populism, who knows that large reforms that eliminate it are not feasible, and desires to know the effects of small reforms on how frequently the voter drains the swamp (the frequency of populist cycles).[12] We divide the analysis in two cases: informed and preemptive populism. Therefore our results inform the choice of a reformer as follows. The reformer may have observed that the voter is draining the swamp sometime against his own interest (for example, the reformer may know that the agent has detected a threat and yet the voter drained the swamp). In this case the reformer should confidently apply Proposition 4. Otherwise, the reformer may be convinced that the voter is only draining the swamp when informed that the agent has violated her mandate. In this case the reformer should only apply Proposition 3. Finally, if the reformer is unsure, a prudent approach is to seek for policies that reduce the frequency of populist cycles across the two propositions.

We begin by stating our results concerning informed populism.

Proposition 3 (Managing informed populism.)**.** *In informed populism, marginal increases in transparency, τ, or the agent's competence, κ, increase the frequency with which the voter drains the swamp. There is no marginal effect of changes in the agent's hawkishness, α.*

In informed populism, the agent always triggers the emergency policy and the voter drains the swamp if and only if he becomes informed that the agent violated her mandate. As transparency increases, the voter becomes informed more often, therefore increasing the frequency with which she drains the swamp. Because emergencies are unlikely ($\pi < 1/2$), the unconditional probability that the agent detects a threat decreases with the agent's competence. But the agent always triggers the emergency policy anyway. So, with a more competent agent, the probability that the voter becomes informed that the agent violated her mandate (i.e., the agent triggered an emergency policy without having detected a threat) increases, increasing the frequency with which the voter drains the swamp. Finally, the agent's hawkishness affects neither the agent's probability of detecting a threat nor the probability that the voter becomes informed, and therefore has no effect on the frequency of informed populist cycles.

Proposition 4 states our results concerning preemptive populism.

Proposition 4 (Managing preemptive populism.)**.** *In preemptive populism, a marginal increase in the agent's hawkishness, α, increases the frequency with which the voter drains the swamp; a marginal increase in transparency, τ, decreases it; and a marginal increase in the agent's competence, κ, may both increase or decrease it, or even have a non-monotonic effect.*

[12]Our approach is aligned with robustness concerns. Our focus on marginal reforms implies that the effects we uncover hold at the margin *and* for larger reforms; in contrast, the effect of large reforms may not hold if the size of the reform is short of what is necessary to eliminate populism altogether. In fact, as we show in Propositions 3 and 4, attempting large reforms that would eliminate populism if large enough may increase the frequency of populist cycles if too small.

In preemptive populism, the voter drains the swamp both when he becomes informed that the agent has violated her mandate and sometime preemptively. A more hawkish agent is more inclined to violate her mandate, thus intensifying her disagreement with the voter. Therefore, disciplining her requires the voter to preemptively drain the swamp more frequently. In contrast, transparency affects the probability with which the voter drains the swamp through two effects. Mechanically, more transparency increases the probability that the voter will become informed that the agent has violated her mandate—and, as a consequence, drain the swamp. However, this also implies that the voter's informed choice to drain the swamp is a more powerful disciplining tool for the agent. This reduces the frequency with which the voter needs to drain the swamp preemptively. Proposition 4 says that this strategic effect dominates the mechanical one.

The marginal impact of a more competent agent is a combination of three distinct effects. A more competent agent detects threats less often, as emergencies are unlikely. Mechanically, keeping all strategies constant, the agent is then triggering the emergency policy more rarely, thus reducing the frequency with which the voter drains the swamp. Furthermore, because disagreement is now less intense, there is less need to discipline the agent: as the agent is violating her mandate less often, the voter would prefer to preemptively drain the swamp less, thus further reducing the frequency of populist cycles. However, such a reduction in discipline would induce the agent to increase the frequency with which she violates her mandate, thus—all else equal—increasing the frequency of populist cycles. Proposition 4 says that the cumulative effect of these three forces can both result in an increase or in a decrease in the frequency of populist cycles, depending on the agent's competence, κ, itself, and on the value of other parameters.

Returning to the reformer's problem, our results can be summarized as follows. If the reformer has evidence that the country has plunged into informed populism, our results recommend to—perhaps counterintuitively—reduce transparency (for example, by decreasing the resources afforded to oversight agencies) or reduce the average competence of experienced public servants. If instead the reformer has evidence that the country is experiencing cycles of preemptive populism, then the reformer should try to increase transparency or decrease disagreement by reducing the hawkishness of public servants (for example, by reducing labor protection differences between private and public sectors). In this case, reforms that attempt to manipulate disagreement through changes in the competence of public servants may backfire if their marginal effect cannot be carefully calculated. Finally, if the reformer is unsure as to whether the country is in informed of preemptive populism, the only prudent reform is to reduce the hawkishness of public servants.

The reformer has indeed one further lever that could be pulled to affect the frequency of populist cycles, and also to eliminate populism altogether: increasing the power of public servants by means of increasing their effectiveness, η. This could be achieved by, for example, increasing the resources available to them in the production of public goods. Proposition 5 informs us in two ways about the result of such a reform. First, it says that marginal increases in the agent's effectiveness in preemptive populism result in more frequent cycles of preemptive populism. Second, in informed populism, increases in effectiveness have no effect, until they indeed produce the disappearance of populism. However, rather than yielding a responsive democracy in which public servants abide by their mandate to serve the voters, such a reform leads to a technocracy in which powerful public servants are not disciplined at all and are able to dictate policy.

Proposition 5 (Managing effectiveness.)**.** *In preemptive populism, a marginal increase in the agent's effectiveness, η, increases the frequency with which the voter drains the swamp. In informed populism, it has no effect on the frequency with which the voter drains the swamp until it induces a technocracy.*

Intuitively, in preemptive populism, greater effectiveness begets greater power, so that the voter is less willing to drain the swamp and replace an experienced agent with a novice, mechanically decreasing the frequency of populist cycles. However, this enables the agent to violate their mandate more often, therefore triggering the emergency policy with greater frequency, causing the voter to drain the swamp more often—more frequent cycles. The first part of Proposition 5 says that in fact in equilibrium this strategic effect dominates so that an increase in effectiveness brings about more frequent cycles of populism.

Returning to the reformer's problem, the result above says that when in a regime of preemptive populism, the reformer would benefit from less effective public servants. On the contrary, in informed populism, the reformer could consider increasing the effectiveness of public servants. However, only large reforms would have any effect and they would lead to technocracy rather than a well-functioning democracy.

7 Discussion of the model

Our benchmark model makes a few simplifications that greatly improve tractability and sharpen the analysis. We now discuss how relaxing these simplifications changes (or does not change) the strategic incentives and equilibrium behavior.

Forward-looking agent. Our benchmark model assumes that an experienced agent receives a payoff V for being retained (or, equivalently, suffers a cost V when the voter drains the swamp), but the agent is not forward-looking in the sense that she does not care about future policies even if retained. This is a natural assumption in our main application if periods are understood to be sufficiently long. In fact, experienced public servants are routinely replaced by new cohorts in the next period. This turnover does not compromise the effectiveness and competence of the organization as the exiting cohort trains and transmits know-how to the new one. Therefore, when voters do not drain the swamp, they retain an experienced organization, but the current members of the public service do not individually participate in future decisions. In contrast, draining the swamp prematurely dismisses the current public service leadership, interrupting the process of transmission of know-how, and causing both a destruction of organizational experience and a personal cost for the current members of the public service.

However, in other applications, or if the periods in our model are supposed to be shorter, it is more reasonable to think that the agent values future policy decisions she may be able to take in the future. We can extend our analysis to incorporate agents that are forward-looking and (potentially) infinitely-lived, maintaining the assumption that, if the voter drains the swamp, the incumbent agent ceases to live and is immediately replaced by a new agent. In this setup, the agent's continuation payoff from the benchmark model, V, is endogenously determined by

$$V(\sigma) := \mathbb{E}\left[\sum_{t'=t+1}^{\infty} \delta^{t'-t-1} \mathbb{I}_{\left(d_{\tilde{t}}=1 \ \forall \tilde{t} \in \{t+1,\ldots,t'\}\right)} \left(\mathbb{I}_{i_{t'}=\theta_{t'}=1} + (1-\alpha)\mathbb{I}_{i_{t'}=\theta_{t'}=0}\right) \,\Big|\, \sigma\right].$$

As per the benchmark model, the agent obtains continuation payoff $V(\sigma)$ if the voter does not drain the swamp; otherwise, they obtain payoff zero. Appendix D characterizes equilibrium behavior in this extended model.

Allowing for a forward-looking agent introduces a further strategic tradeoff to the ones in our benchmark model. In fact, in the extended model, when the voter attempts to discipline the agent by draining the swamp more often, he needs to take into account that now draining the swamp is a double-edged sword. On the one hand, draining the swamp reduces the agent's expected payoff when she triggers the emergency policy. This has a disciplining effect. However, now there is also a dynamic effect: draining the swamp more often reduces the agent's expected payoff of being retained, thus reducing the benefit of abiding by her mandate. While this new tradeoff complicates the analysis significantly, we show in Appendix D that our results carry on under a technical assumption that guarantees a unique solution to the voter's problem in preemptive populism.[13]

Multi-period accumulation of experience ($T > 1$). In our benchmark model, a novice hired in period t becomes experienced in period $t + T = t + 1$. If the periods in our model are sufficiently short, it may be more reasonable to assume that a new agent may need $T > 1$ periods before she accumulates sufficient experience to enhance her effectiveness and competence.

Appendix E studies this extended model. We show that, as in our benchmark model, the incompetence of experienced agents spurs disagreement between them and the voter, and that their effectiveness begets power. As a consequence, the main results of Proposition 1 continue to hold under a technical assumption that guarantees a unique solution to the voter's problem in preemptive populism. As in our benchmark model, a sufficiently powerful agent establishes a technocracy, and a responsive democracy requires less intense disagreement. When disagreement does not allow for a responsive democracy, populism arises, and a less powerful agent will induce the voter to preemptively drain the swamp.

However, the lapse of time required for the agent to accumulate experience is not inconsequential. In fact, when $T > 1$, draining the swamp is more costly for the voter, as he knows that replacing an experienced agent with a novice induces a cost for several periods. Importantly, this cost of draining the swamp is greater if experienced agents abide by their mandate. Intuitively, if the voter expects an experienced agent to abide by her mandate, then he knows that she will provide both more public goods and better policies than a novice. In contrast, if he expects the agent to violate her mandate, then the benefit of an experienced agent is limited to the provision of public goods and, in fact, experienced agents provide worse policies, on average, than novices. This implies that the power of the agent increases with T.

This increase in the power of the agent makes technocracy more likely and preemptive populism and responsive democracy less likely. Furthermore, a new form of power-induced informed populism may arise. In fact, for some parameters, the agent is at the same time sufficiently powerful so that the voter is unable to fully discipline her and yet not so powerful that she can establish a technocracy. In equilibrium the agent sometimes violates her mandate and the voter, when he becomes informed of the violation, sometimes, but not always, drains the swamp.

Proposition 6 (Power-induced informed populism). *If $T > 1$, there exists $H \subset \mathbb{R}_{++}$ such*

[13]This assumption is also satisfied at the limit when δ is arbitrarily small so that our benchmark and extended model coincide. When the technical assumption is not satisfied, the model is analytically intractable.

that, if $\eta \in H$ and $\underline{\kappa}(\pi, \alpha, \tau) < \kappa$, then, in every period, the agent violates her mandate with strictly positive probability and the voter drains the swamp with probability strictly between 0 and 1 if he becomes informed that the agent violated her mandate and does not drain the swamp otherwise.

In this regime, draining the swamp occurs with positive probability, but only if the voter is informed. However, the agent is so powerful that sometimes the voter does not drain the swamp even when informed that the agent violated her mandate.

From a policy prospective, this means that, when experience takes longer to accumulate, informed anti-elite populism (but not preemptive populism) may also arise as a consequence of public servants becoming more effective.

8 Conclusion

We have studied a theoretical framework that can rationally explain why voters may want to replace experienced public servants with less effective and less competent novices. In equilibrium, when the state bureaucracy is not too powerful and disagrees enough with the voters, draining the swamp always arises. But more power for the bureaucracy does not deliver a more responsive state. Instead, it yields a technocracy in which bureaucrats govern for themselves rather than to serve the voters. A responsive democracy, in which the state works for the citizens and the citizens trust the state only works in a "narrow corridor" (Acemoglu and Robinson, 2019) where the public servants' disagreement with the voters and their power are both moderate.

We showed that our theory offers insights into which technological and political shocks can intensify disagreement between voters and public servants and weaken the implicit power of public servants, thus causing voters to drain the swamp. Some of these shocks may be irreversible even for skillful reformers. Furthermore, some reform levers may backfire—increasing, instead of decreasing, the frequency with which voters drain the swamp—if not well calibrated.

The only reform that unambiguously reduces voters' inclination to drain the swamp is one that reduces disagreement in the most direct way: it reduces the differences in preferences between bureaucrats and voters. A more inclusive and representative state bureaucracy thus is the best guarantee against populism, as it induces voters to trust the state more and reduces their need to "control" the state by draining the swamp. This result echoes arguments that link the historical development of merit-based bureaucracies in China with greater trust in the central state. However, as our model warns us, even if representative, a too effective state is also one that can make democracy unsustainable and establish a powerful technocracy by which the state governs for itself (see, e.g., Stasavage, 2020).

References

Acemoglu, Daron and James A Robinson, *The narrow corridor: How nations struggle for liberty*, Penguin UK, 2019.

Alesina, Alberto and Guido Tabellini, "Bureaucrats or politicians? Part I: a single policy task," *American Economic Review*, 2007, *97* (1), 169–179.

Alesina, Alberto and Guido Tabellini, "Bureaucrats or politicians? Part II: Multiple policy tasks," *Journal of Public Economics*, 2008, *92* (3-4), 426–447.

Algan, Yann, Sergei Guriev, Elias Papaioannou, and Evgenia Passari, "The European trust crisis and the rise of populism," *Brookings Papers on Economic Activity*, 2017, *2017* (2), 309–400.

Ananyev, Maxim and Sergei Guriev, "Effect of income on trust: evidence from the 2009 economic crisis in Russia," *The Economic Journal*, 2019, *129* (619), 1082–1118.

Ashworth, Scott and Ethan Bueno de Mesquita, "Is voter competence good for voters?: Information, rationality, and democratic performance," *American Political Science Review*, 2014, *108* (3), 565–587.

Ashworth, Scott, Ethan Bueno de Mesquita, and Amanda Friedenberg, "Learning about voter rationality," *American Journal of Political Science*, 2018, *62* (1), 37–54.

Autor, David, David Dorn, Gordon Hanson, and Kaveh Majlesi, "Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure," *American Economic Review*, 2020, *110* (10), 3139–3183.

Banks, Jeffrey S and Barry R Weingast, "The political control of bureaucracies under asymmetric information," *American Journal of Political Science*, 1992, *36* (2), 509–524.

Banks, Jeffrey S and Joel Sobel, "Equilibrium selection in signaling games," *Econometrica*, 1987, pp. 647–661.

Barro, Robert J, "The control of politicians: an economic model," *Public Choice*, 1973, pp. 19–42.

Bauer, Michael W, B Guy Peters, Pierre Jon, Kutsal Yesilkagit, and Stefan Becker, *Introduction: Populists, democratic backsliding, and public administration*, Cambridge University Press, 2021.

Bellodi, Luca, Massimo Morelli, and Matia Vannoni, "A Costly Commitment: Populism, Government Performance, and the Quality of Bureaucracy," *American Journal of Political Science*, forthcoming.

Bendor, Jonathan and Adam Meirowitz, "Spatial models of delegation," *American Political Science Review*, 2004, *98* (2), 293–310.

Besley, Timothy, *Principled agents?: The political economy of good government*, Oxford University Press on Demand, 2006.

Bovard, Rachel, "The Trump administration's successful war against bureaucratic bullies," *USA Today*, 2019, (October 16).

Calabresi, Massimo, "Inside Donald Trump's War Against the State," *Time Magazine*, 2017, (March 20).

Callander, Steven and Keith Krehbiel, "Gridlock and delegation in a changing world," *American Journal of Political Science*, 2014, *58* (4), 819–834.

Cho, In-Koo and David M Kreps, "Signaling games and stable equilibria," *The Quarterly Journal of Economics*, 1987, *102* (2), 179–221.

Colantone, Italo and Piero Stanig, "The Trade Origins of Economic Nationalism: Import Competition and Voting Behavior in Western Europe," *American Journal of Political Science*, 2018, *62* (4), 936–953.

Costa, Robert, "Ben Carson is unlikely to serve in Trump's Cabinet," *Washington Post*, 2016, (November 15).

Dewan, Torun and Francesco Squintani, "Leadership with trustworthy associates," *American Political Science Review*, 2018, *112* (4), 844–859.

Duggan, John and César Martinelli, "The political economy of dynamic elections: Accountability, commitment, and responsiveness," *Journal of Economic Literature*, 2017, *55* (3), 916–984.

Ferejohn, John, "Incumbent performance and electoral control," *Public Choice*, 1986, pp. 5–25.

Fox, Justin and Kenneth W Shotts, "Delegates or trustees? A theory of political accountability," *The Journal of Politics*, 2009, *71* (4), 1225–1237.

Fox, Justin and Stuart V Jordan, "Delegation and accountability," *The Journal of Politics*, 2011, *73* (3), 831–844.

Fudenberg, Drew and Jean Tirole, "Perfect Bayesian Equilibrium and Sequential Equilibrium," *Journal of Economic Theory*, 1991, *53* (2), 236–260.

Gailmard, Sean and John W Patty, "Formal models of bureaucracy," *Annual Review of Political Science*, 2012, *15*, 353–377.

Gailmard, Sean and John W Patty, *Learning while governing: Expertise and accountability in the executive branch*, University of Chicago Press, 2012.

Gratton, Gabriele and Barton E Lee, "Liberty, security, and accountability: The rise and fall of illiberal democracies," *The Review of Economic Studies*, forthcoming.

Gratton, Gabriele, Luigi Guiso, Claudio Michelacci, and Massimo Morelli, "From Weber to Kafka: Political instability and the overproduction of laws," *American Economic Review*, 2021, *111* (9), 2964–3003.

Guiso, Luigi, Helios Herrera, Massimo Morelli, and Tommaso Sonno, "Populism: Demand and Supply," 2017.

Guiso, Luigi, Helios Herrera, Massimo Morelli, and Tommaso Sonno, "Global Crises and Populism: The Role of Eurozone Institutions," *Economic Policy*, 2019, *34* (97), 95–139.

Guiso, Luigi, Helios Herrera, Massimo Morelli, and Tommaso Sonno, "Economic Insecurity and the Demand of Populism in Europe," *Einaudi Institute for Economics and Finance Working Paper*, 2020.

Guriev, Sergei and Elias Papaioannou, "The Political Economy of Populism," *CEPR Discussion Paper No. DP14433*, 2020.

Holmström, Bengt, "On the theory of delegation," *Northwestern University, Center for Mathematical Studies in Economics and Management Science Discussion Paper No. 438*, 1980.

Holmström, Bengt, "Managerial incentive problems: A dynamic perspective," *Essays in Economic and Management in Honor of Lars Wahbleck. Reprinted in The Review of Economic Studies 1999*, 1982, *66*, 169–182.

Huber, John D and Charles R Shipan, *Deliberate discretion?: The institutional foundations of bureaucratic autonomy*, Cambridge University Press, 2002.

Huber, John D and Nolan McCarty, "Bureaucratic capacity, delegation, and political reform," *American Political Science Review*, 2004, *98* (3), 481–494.

Kartik, Navin, Richard Van Weelden, and Stephane Wolton, "Electoral ambiguity and political representation," *American Journal of Political Science*, 2017, *61* (4), 958–970.

Maskin, Eric and Jean Tirole, "The principal-agent relationship with an informed principal, II: Common values," *Econometrica*, 1992, *60* (1), 1–42.

Maskin, Eric and Jean Tirole, "The politician and the judge: Accountability in government," *American Economic Review*, 2004, *94* (4), 1034–1054.

Morelli, Massimo, Antonio Nicolò, and Paolo Roberti, "A commitment theory of populism," 2021.

Morris, Stephen, "Political correctness," *Journal of Political Economy*, 2001, *109* (2), 231–265.

Moynihan, Donald, *Populism and the Deep State: The Attack on Public Service Under Trump*, Cambridge University Press,

Mudde, Cas and Cristóbal Rovira Kaltwasser, *Populism: A Very Short Introduction*, Oxford University Press, 2017.

Nath, Anousha et al., "Bureaucrats and politicians: Electoral competition and dynamic incentives," *Unpublished working paper. Boston University*, 2016.

Norris, Pippa and Ronald Inglehart, *Cultural backlash: Trump, Brexit, and authoritarian populism*, Cambridge University Press, 2019.

Panizza, Francisco, B Guy Peters, and Conrado R Ramos Larraburu, "Roles, trust and skills: A typology of patronage appointments," *Public Administration*, 2019, *97* (1), 147–161.

Postel, Charles, *The populist vision*, Oxford University Press, 2007.

Przeworski, Adam, *Crises of democracy*, Cambridge University Press, 2019.

Rodrik, Dani, "Is populism necessarily bad economics?," in "AEA Papers and proceedings," Vol. 108 2018, pp. 196–99.

Sasso, Greg and Massimo Morelli, "Bureaucrats under populism," *Journal of Public Economics*, 2021, *202*, 104497.

Stasavage, David, *The Decline and Rise of Democracy: A Global History from Antiquity to Today*, Vol. 96, Princeton University Press, 2020.

Ting, Michael M, "A theory of jurisdictional assignments in bureaucracies," *American Journal of Political Science*, 2002, pp. 364–378.

Ting, Michael M, "The political economy of governance quality," *American Political Science Review*, 2021, *115* (2), 667–685.

Appendix

A Equilibrium definition

We study the unique perfect Bayesian equilibrium (Fudenberg and Tirole, 1991) in Markovian strategies that survives divinity. A Markovian strategy for the agent is a mapping p from the signal realization, $s_t \in \{0,1\}$, to a probability of triggering the emergency policy. A Markovian strategy for the voter is a mapping d from the observed signal (or lack thereof), $\hat{s}_t \in \{\emptyset, s_t\}$, to a probability of draining the swamp (if the agent has triggered the emergency policy). We denote by μ_t the agent's belief that $\theta_t = 1$—a mapping from the signal realization, s_t, into a probability. We denote by ν_t the voter's belief that $\theta_t = 1$ when the agent has triggered the emergency policy and upon observing the signal (or lack thereof)—a mapping from the observed signal, $\hat{s}_t$, into a probability. A Markovian assessment is a tuple $\sigma = (p, d, \{\mu_t\}_{t=1}^{\infty}, \{\nu_t\}_{t=1}^{\infty})$. Let

$$U\left(d_t \mid \sigma\right) := \mathbb{E}\left[\sum_{t'=t+1}^{\infty} \delta^{t'-t-1}\left(\mathbb{I}_{i_{t'}=\theta_{t'}} + \eta \mathbb{I}_{d_{t'}=0}\right) \,\middle|\, \sigma\right]$$

denote the voter's expected continuation payoff from $d_t \in \{0,1\}$, given σ. A strategy p^* for the agent is sequential rational, given σ, if it maximizes the agent's expected payoff in each period t and for each signal realization s_t, i.e.,

$$p^*(s_t) \in \arg\max_{p\in[0,1]} \left\{ \begin{array}{c} p\left[\tau\left(1-d(s_t)\right)+(1-\tau)\left(1-d(\emptyset)\right)\right]\left(\mu_t(s_t)+\delta V\right) \\ +(1-p)\left[(1-\alpha)(1-\mu_t(s_t))+\delta V\right] \end{array} \right\} \tag{A.1}$$

A strategy d^* for the voter is sequentially rational, given σ, if it maximizes the voter's expected payoff in each period t and for each observed signal (or lack thereof) $\hat{s}_t$, , i.e.,

$$d^*(\hat{s}_t) \in \arg\max_{d\in[0,1]} \left\{ \begin{array}{c} d\left[(1-\nu_t(\hat{s}_t))+\delta U(d_t=1\mid\sigma)\right] \\ +(1-d)\left[\nu_t(\hat{s}_t)+\eta+\delta U(d_t=0\mid\sigma)\right] \end{array} \right\} \tag{A.2}$$

We now discuss which beliefs are consistent with σ. The agent's beliefs are always uniquely pinned down by Bayes' rule:

$$\mu_t(0) = \mu_t^*(0) := \Pr[\theta_t = 1 \mid s_t = 0] = \frac{\pi(1-\kappa)}{\pi(1-\kappa)+(1-\pi)\kappa} \tag{A.3}$$

and

$$\mu_t(1) = \mu_t^*(1) : = \Pr[\theta_t = 1 \mid s_t = 1] = \frac{\pi\kappa}{\pi\kappa + (1-\pi)(1-\kappa)}. \tag{A.4}$$

Notice that consistency requires that the voter's beliefs are bounded: for any $\hat{s}_t$,

$$\nu_t(\hat{s}_t) \in [\nu_t^*(0), \nu_t^*(1)] \tag{A.5}$$

because the agent's action cannot signal information that she does not possess.[14] Furthermore, the voter's beliefs are also uniquely pinned down by Bayes' rule whenever she directly observes the agent's signal:[15]

$$\nu_t(0) = \nu_t^*(0) : = \mu_t^*(0) \qquad \text{and} \qquad \nu_t(1) = \nu_t^*(1) := \mu_t^*(1); \tag{A.6}$$

and otherwise whenever the agent triggers the emergency policy with strictly positive probability (i.e., $p^*(s_t) > 0$ for some s_t):

$$\begin{aligned} \nu_t(\emptyset) = \nu_t^*(\emptyset) &: = \Pr[\theta_t = 1 \mid p_t = 1, \sigma, \hat{s}_t = \emptyset] \\ &= \frac{\pi\kappa p^*(1) + \pi(1-\kappa)p^*(0)}{(\pi\kappa + (1-\pi)(1-\kappa))p^*(1) + (\pi(1-\kappa) + (1-\pi)\kappa)p^*(0)}. \end{aligned} \tag{A.7}$$

We impose a further condition on the voter's belief $\nu_t^*(\emptyset)$ when the voter does not observe the signal and, in equilibrium, the agent never triggers the emergency policy. Adopting Cho and Kreps (1987)'s definition (see, e.g., Fudenberg and Tirole, 1991; Maskin and Tirole, 1992), we say that an equilibrium is divine if $\nu_t^*(\emptyset)$ satisfies condition D1. Formally, given sequentially rational $d^*(0)$ and $d^*(1)$, let $d_{\emptyset,s}$ be the value of $d(\emptyset)$ such that, upon observing $s_t = s$, the agent is indifferent between $p_t = 1$ and $p_t = 0$; that is,

$$[\tau d^*(s) + (1-\tau)d_{\emptyset,s}]\,(\mu_t^*(s) + \delta V) = (1-\alpha)(1-\mu_t^*(s)) + \delta V.$$

In our context, condition D1 says that, whenever $p^*(s_t) = 0$ for all s_t, if there exists s and s' such that $d_{\emptyset,s'} < d_{\emptyset,s}$, then $\nu_t^*(\emptyset) = \nu_t^*(s')$.

Definition A.1 (Equilibrium)**.** *An assessment $\sigma = (p^*, d^*, \{\mu_t^*\}_{t=1}^\infty, \{\nu_t^*\}_{t=1}^\infty)$ is an equilibrium if, for each t, p^* satisfies* (A.1) *for $\sigma = \sigma^*$; d^* satisfies* (A.2) *for $\sigma = \sigma^*$; $\mu_t^*(s_t)$ satisfies* (A.3)–(A.4)*; and $\nu_t^*(\hat{s}_t)$ satisfies* (A.5)–(A.6), (A.7) *if well-defined, and condition D1.*

[14] This is essentially a version of the "no signaling what you don't know" condition (see, e.g., Fudenberg and Tirole, 1991).

[15] Notice that this is true also off the equilibrium path.

B Omitted proofs

Proof of Lemma 1. For any signal $s_t \in \{0, 1\}$, the probability of an emergency is given by $\mu_t^*(0)$ (A.3) and $\mu_t^*(1)$ (A.4). Because $\kappa > 1 - \pi > \pi$, $\mu_t^*(0) < 1/2 < \mu_t^*(1)$. Therefore, all else equal, the voter prefers $i_t = 1$ if and only if $s_t = 1$. In contrast—and holding all else equal—the agent prefers $i_t = 1$ if and only if $\mu_t^*(\hat{s}) > (1 - \mu_t^*(\hat{s}))(1 - \alpha)$, which yields $\mu_t^*(\hat{s}) > (1 - \alpha)/[1 + (1 - \alpha)] \in (0, 1/2)$ where $\hat{s} \in \{0, 1\}$ is the agent's signal. Because $\mu_t^*(1) > 1/2$, the last inequality holds for $\hat{s} = 1$; it holds for $\hat{s} = 0$ if and only if $\kappa < \bar{\kappa}(\alpha, \pi) := \pi/[\pi + (1 - \alpha)(1 - \pi)]$. $\square$

Proof of Lemma 2. Suppose $p_t = 1$. The voter's expected payoff from $d_t = 1$ is $(1 - \nu_t) + \delta U(\sigma^*)$; his expected payoff from $d_t = 0$ is $\nu_t + \eta + \delta U(\sigma^*)$. Therefore, he chooses $d_t = 1$ if and only if $\nu_t < \frac{1-\eta}{2}$. $\square$

Proof of Lemma 3. Suppose $p_t = 1$. If the voter observes $s_t = 1$. Then, by Bayes' rule (see Definition A.1), $\nu_t(s_t) > 1/2$ and, by Lemma 2, he chooses $d_t = 0$. If instead the voter observes $s_t = 0$, then, by Bayes' rule (see Definition A.1) and Lemma 2, the voter chooses $d_t = 1$ if and only if

$$\frac{\pi(1 - \kappa)}{\pi(1 - \kappa) + (1 - \pi)\kappa} < \frac{1 - \eta}{2} \iff \eta < \bar{\eta}(\pi, \kappa) := 1 - 2\frac{\pi(1 - \kappa)}{\pi(1 - \kappa) + (1 - \pi)\kappa}. \qquad \square$$

Proof of Lemma 4. Let $\eta \geq \bar{\eta}(\pi, \kappa)$. By Lemma 3, the voter chooses $d_t = 0$ when informed that $s_t = 0$ or $s_t = 1$. When uninformed, by Definition A.1, the voter's belief is contained in the interval $[\nu_t(s_t = 0), \nu_t(s_t = 1)]$ (see (A.5)) and, hence, also chooses $d_t = 0$. Therefore, by Assumption 2, the agent optimally chooses $p_t = 1$ for all s_t.

Let $\eta < \bar{\eta}(\pi, \kappa)$. By Lemma 3, in every equilibrium, the voter chooses $d_t = 1$ when informed if and only if $s_t = 0$. We now establish a sequence of useful auxiliary lemmas.

Lemma B.1. *Let $\eta < \bar{\eta}(\pi, \kappa)$. In every equilibrium, $p^*(0) \leq p^*(1)$.*

Proof. For sake of a contradiction, suppose $p^*(1) < p^*(0)$ and, hence, $0 < p^*(0)$ and $p^*(1) < 1$. Because $p^*(0) > 0$ is optimal for the agent, then

$$(1 - \tau)d^*(\mu_t^*(0) + \delta V) \geq (1 - \alpha)(1 - \mu_t^*(0)) + \delta V. \tag{B.1}$$

Similarly, because $p^*(1) < 1$ is optimal for the agent, then

$$(1-\alpha)(1-\mu_t^*(1)) + \delta V \geq (\tau + (1-\tau)d^*)(\mu_t^*(1) + \delta V). \tag{B.2}$$

Combining (B.1) and (B.2) we obtain

$$(1-\tau)d^*(\mu_t^*(0) + \delta V) \geq (\tau + (1-\tau)d^*)(\mu_t^*(1) + \delta V), \tag{B.3}$$

which, because $\mu_t^*(0) < \mu_t^*(1)$, is a contradiction. □

Lemma B.2. *Suppose $\eta < \bar{\eta}(\pi, \kappa)$. There is no equilibrium with $p^*(0) \in (0,1)$ and $p^*(1) \in (0,1)$.*

Proof. Similar to the proof of Lemma B.1: Inequalities (B.1) and (B.2), implies the same contradiction (i.e., Inequality (B.3)). □

Lemma B.3. *Let $\eta < \bar{\eta}(\pi, \kappa)$. If in equilibrium $p^*(0) = 0$ and $p^*(1) > 0$, then $p^*(1) = 1$.*

Proof. For sake of a contradiction, suppose $p^*(0) = 0$ and $p^*(1) \in (0,1)$. By Definition A.1, $\nu_t^*(\emptyset) = \nu_t^*(1) = \mu_t^*(1)$ and therefore $d^* = 0$. Then, when $s_t = 1$ the agent's expected payoff from $p_t = 1$ is $\mu_t^*(1) + \delta V > (1-\alpha)(1-\mu^*(1)) + \delta V$, which is their expected payoff from $p_t = 0$. Therefore, the agent optimally chooses $p^*(1) = 1$—a contradiction. □

Lemma B.4. *Let $\eta < \bar{\eta}(\pi, \kappa)$. There is no equilibrium with $p^*(0) = p^*(1) = 0$.*

Proof. For sake of a contradiction, suppose $p^*(0) = p^*(1) = 0$. The voter's belief when uninformed, $\nu_t^*(\emptyset)$, is off the equilibrium path. We apply the D1 condition: Suppose $d^* = d'$ for some d'. When $s_t = 0$, the agent will be indifferent between $p_t = 1$ and $p_t = 0$ if and only if

$$(1-\alpha)(1-\mu^*(0)) + \delta V = (1-\tau)d'(\mu^*(0) + \delta V). \tag{B.4}$$

When $s_t = 1$, the agent will be indifferent if and only if

$$(1-\alpha)(1-\mu^*(1)) + \delta V = (\tau + (1-\tau)d')(\mu^*(1) + \delta V). \tag{B.5}$$

Because $\mu^*(1) > \mu^*(0)$, the value of d' such that (B.4) holds is strictly greater than the value of d' such that (B.5) holds. Thus, the D1 condition requires $\nu_t^*(\emptyset) = \nu_t^*(s_t = 1)$. But then, by Lemma 2, $d^* = 0$, which immediately leads to a contradiction: when $s_t = 1$, the agent strictly prefers to choose $p_t = 1$. □

Together Lemmas B.1–B.4 imply that, in equilibrium, one of three must hold: (i) $p^*(0) = 0$ and $p^*(1) = 1$; (ii) $p^*(0) \in (0,1)$ and $p^*(1) = 1$; or (iii) $p^*(0) = 1$ and $p^*(1) = 1$. We can then conclude that in every equilibrium, the agent chooses $p_t = 1$ when $s_t = 1$ and, by sequential rationality, when $s_t = 0$, the agent chooses $p_t = 1$ with strictly positive probability only if

$$(1-\alpha)(1-\mu^*(0)) + \delta V \leq (1-\tau)d^*(\mu^*(0) + \delta V). \qquad \square$$

Proof of Proposition 1. **Part (i)** follows immediately from Lemmas 2–4. **For Parts (ii)–(iv)**, we characterize the set of parameters for which each regimes exists. The interior of each set of parameters will be distinct and, hence, it follows that the equilibrium is unique with the exception of a measure-zero set of parameters. Let $\eta < \bar{\eta}(\pi, \kappa)$. By Lemmas 3 and 4, $p^*(1) = 1$, $d^*(0) = 1$, and $d^*(1) = 1$.

Part (ii): In a responsive democracy, $d^*(\emptyset) = 0$ and $p^*(0) = 0$. By Definition A.1, $\nu_t^*(\emptyset) = \nu_t^*(1)$, so that $d^*(\emptyset) = 0$. $p^*(0) = 0$ is optimal if and only if

$$(1-\alpha)(1-\mu_t^*(0)) + \delta V \geq (1-\tau)(\mu_t^*(0) + \delta V). \tag{B.6}$$

Both sides of (B.6) are continuous in κ and the left (resp., right) hand side is increasing (resp., decreasing) in κ. Furthermore, at $\kappa = \bar{\kappa}(\pi, \alpha)$ (see Proof of Lemma 1), we have

$$(1-\alpha)(1-\mu(0)) = \mu(0) \implies (1-\alpha)(1-\mu(0)) + \delta V > (1-\tau)(\mu(0) + \delta V).$$

Define $\underline{\kappa}(\pi, \alpha, \tau) \in (1-\pi, \bar{\kappa}(\pi, \alpha))$ as the value of κ for which (B.6) holds with equality or, if such a value does not exist, $\underline{\kappa}(\pi, \alpha, \tau) = 1 - \pi$. Therefore, a responsive democracy is an equilibrium if and only if $\eta < \bar{\eta}(\pi, \kappa)$ and $\underline{\kappa}(\pi, \alpha, \tau) \leq \kappa$.

Part (iii): In informed populism, $d^*(\emptyset) = 0$ and $p^*(0) = 1$. By Definition A.1, $\nu_t^*(\emptyset) = \pi$ so that, by Lemma 2, $d^*(\emptyset) = 0$ is optimal if and only if

$$\pi \geq \frac{1-\eta}{2} \iff \eta \geq \underline{\eta}(\pi) := 1 - 2\pi. \tag{B.7}$$

$p^*(0) = 1$ is optimal if and only if

$$(1-\alpha)(1-\mu(0)) + \delta V \leq (1-\tau)(\mu(0) + \delta V), \tag{B.8}$$

which is the reversed inequality of (B.6) and therefore, for $\kappa > 1 - \pi$, (B.8) holds if and only if $\kappa \leq \underline{\kappa}(\pi, \alpha, \tau)$. Thus, informed populism is an equilibrium if and only if $\underline{\eta}(\pi) \leq$

$\eta \leq \bar{\eta}(\pi, \kappa)$ and $\kappa \leq \underline{\kappa}(\pi, \alpha, \tau)$.

Part (iv): In preemptive populism, $d^*(\emptyset) \in (0,1)$ and $p^*(0) \in (0,1)$. By Definition A.1,

$$\nu_t^*(s_t = \emptyset) = \frac{\pi\kappa + \pi(1-\kappa)p^*}{\Pr[s_t = 1] + \Pr[s_t = 0]p^*},$$

so that, by Lemma 2, $d^*(\emptyset) \in (0,1)$ is optimal if and only if

$$\frac{\pi\kappa + \pi(1-\kappa)p^*(0)}{\Pr[s_t = 1] + \Pr[s_t = 0]p^*(0)} = \frac{1-\eta}{2}. \tag{B.9}$$

The left hand side of (B.9) is continuous and decreasing in $p^*(0)$. It ranges from π (at $p^*(0) = 1$) to $\mu_t^*(1) > (1-\eta)/2$ (at $p^*(0) = 0$), where the last inequality follows from $\eta < \bar{\eta}(\pi, \kappa)$. Thus, a (unique) solution $p^*(0) \in (0,1)$ to (B.9) exists if and only if $\pi < (1-\eta)/2$: $\eta < \underline{\eta}(\pi)$.

Furthermore, $p^*(0) \in (0,1)$ is optimal if and only if

$$(1-\alpha)(1-\mu_t^*(0)) + \delta V = (1-\tau)(1-d^*(\emptyset))(\mu_t^*(0) + \delta V). \tag{B.10}$$

The right hand side is continuous and decreasing in $d^*(\emptyset)$. It ranges from $0 < (1-\alpha)(1-\mu_t^*(0)) + \delta V$ (when $d^*(\emptyset) = 1$) to $(1-\tau)(\mu(0) + V)$ (when $d^*(\emptyset) = 0$). Thus, a (unique) solution $d^*(\emptyset) \in (0,1)$ to (B.10) exists if and only if

$$(1-\alpha)(1-\mu(0)) + \delta V < (1-\tau)(\mu(0) + \delta V), \tag{B.11}$$

which is the reverse inequality of (B.6) and therefore, for $\kappa > 1 - \pi$, (B.11) holds if and only if $\kappa < \underline{\kappa}(\pi, \alpha, \tau)$. Therefore, preemptive populism is an equilibrium if and only if $\eta < \underline{\eta}(\pi)$ and $\kappa < \underline{\kappa}(\pi, \alpha, \tau)$. □

Proof of Proposition 2. Let $\kappa, \alpha, \pi, \tau$ induce a responsive democracy:

$$\eta < \bar{\eta}(\pi, \kappa) \text{ and } \underline{\kappa}(\pi, \alpha, \tau) < \kappa < \bar{\kappa}(\pi, \alpha), \tag{B.12}$$

where

$$\underline{\kappa}(\pi, \alpha, \tau) = \frac{\pi((1-\tau) - \tau\delta V)}{\pi((1-\tau) - \tau\delta V) + (1-\pi)((1-\alpha) + \tau\delta V)}. \tag{B.13}$$

Note that, because $\pi < 1/2$, the denominator of (B.13) is positive and, hence,

$$\underline{\kappa}(\pi,\alpha,\tau) > 0 \iff (1-\tau) - \tau\delta V > 0. \tag{B.14}$$

Part (i): Sufficiency: Let $1-\pi < \underline{\kappa}(\pi,\alpha,\tau)$ and $\eta < \bar{\eta}(\pi,\underline{\kappa}(\pi,\alpha,\tau))$ and take $\kappa' = \underline{\kappa}(\pi,\alpha,\tau) - \varepsilon$, where $\varepsilon > 0$. For ε sufficiently small, $\kappa' \in (1-\pi,\kappa)$ and, by continuity of $\bar{\eta}$, $\eta < \bar{\eta}(\pi,\kappa'$. Therefore, κ' induces populism. Necessity: Let $\kappa' \in (1-\pi,\kappa)$ induce (preemptive or informed) populism: $\eta < \bar{\eta}(\pi,\kappa')$ and $\kappa' < \underline{\kappa}(\pi,\alpha,\tau)$. Because $\kappa' > 1-\pi$, the second inequality implies that $1-\pi < \underline{\kappa}(\pi,\alpha,\tau)$. Furthermore, because

$$\frac{\partial\bar{\eta}(\pi,\kappa)}{\partial\kappa} = \frac{2(1-\pi)\pi}{(\kappa(1-\pi)+\pi(1-\kappa))^2} > 0 \tag{B.15}$$

and $\kappa' < \underline{\kappa}(\pi,\alpha,\tau)$, we have that $\eta < \bar{\eta}(\pi,\kappa') < \bar{\eta}(\pi,\underline{\kappa}(\pi,\alpha,\tau))$.

Part (ii): Sufficiency: Let $\kappa < \underline{\kappa}(\pi,1,\tau)$. Take $\alpha' = 1-\varepsilon$, where $\varepsilon > 0$. If ε is sufficiently small, then $\alpha' \in (\alpha,1)$ and, by continuity of $\underline{\kappa}$, $\kappa < \underline{\kappa}(\pi,\alpha',\tau)$. Then, because $\bar{\eta}(\pi,\kappa)$ is independent of α and, by (B.12), $\eta < \bar{\eta}(\pi,\kappa)$, α' induces populism. Necessity: Let $\alpha' \in (\alpha,1)$ induce populism: $\eta < \bar{\eta}(\pi,\kappa)$ and $\kappa < \underline{\kappa}(\pi,\alpha',\tau)$. Because $\kappa > 0$, the second inequality implies that $\underline{\kappa}(\pi,\alpha',\tau) > 0$. Using by (B.14), we then obtain

$$\frac{\partial\underline{\kappa}(\pi,\alpha,\tau)}{\partial\alpha} = -\frac{-(1-\pi)\pi((1-\tau)-\tau\delta V)}{((1-\alpha)(1-\pi)+\pi+\tau(\delta V-\pi(1+2\delta V)))^2} > 0.$$

Thus, $\kappa < \underline{\kappa}(\pi,\alpha',\tau) < \underline{\kappa}(\pi,1,\tau)$.

Part (iii): Sufficiency: Let $\kappa < \underline{\kappa}(\pi_\eta,\alpha,\tau)$. Notice that because (B.12) holds and

$$\frac{\partial\bar{\eta}(\pi,\kappa)}{\partial\pi} = \frac{-2(1-\kappa)\kappa}{(\kappa(1-\pi)+\pi(1-\kappa))^2} < 0, \tag{B.16}$$

the definition of π_η implies that $\pi < \pi_\eta$. Take $\pi' = \pi_\eta - \varepsilon$, where $\varepsilon > 0$. If ε is sufficiently small, then $\pi' \in (\pi,1/2)$ and, by continuity of $\underline{\kappa}$, $\kappa < \underline{\kappa}(\pi',\alpha,\tau)$. Since $\eta \leq \bar{\eta}(\pi_\eta,\kappa) < \bar{\eta}(\pi',\kappa)$, π' induces populism. Necessity: Let $\pi' \in (\pi,1/2)$ induce populism: $\eta < \bar{\eta}(\pi',\kappa)$ and $\kappa < \underline{\kappa}(\pi',\alpha,\tau)$. Because $\kappa > 0$, the second inequality implies that $\underline{\kappa}(\pi,\alpha',\tau) > 0$. Using by (B.14), we then obtain

$$\frac{\partial\underline{\kappa}(\pi,\alpha,\tau)}{\partial\pi} = \frac{((1-\alpha)+\tau\delta V)((1-\tau)-\tau\delta V)}{((1-\alpha)(1-\pi)+\pi+\tau(\delta V-\pi(1+2\delta V)))^2} > 0. \tag{B.17}$$

Because $\bar{\eta}(\pi', \kappa)$ is decreasing in π' and $\eta < \bar{\eta}(\pi', \kappa)$, then $\pi_\eta > \pi'$. Therefore, $\kappa < \underline{\kappa}(\pi', \alpha, \tau) < \underline{\kappa}(\pi_\eta, \alpha, \tau)$.

Part (iv): Sufficiency: Let $\kappa < \underline{\kappa}(\pi, \alpha, 0)$. Take $\tau' = \varepsilon$, where $\varepsilon > 0$. If ε is sufficiently small, then $\tau' \in (0, \tau)$ and, by continuity of $\underline{\kappa}$, $\kappa < \underline{\kappa}(\pi, \alpha, \tau')$. Then, because $\bar{\eta}(\pi, \kappa)$ is independent of τ and, hence, by (B.12), $\eta < \bar{\eta}(\pi, \kappa)$, τ' induces populism. Necessity: Let $\tau' \in (0, \tau)$ induce populism: $\eta < \bar{\eta}(\pi, \kappa)$ and $\kappa < \underline{\kappa}(\pi, \alpha, \tau')$. Because

$$\frac{\partial \underline{\kappa}(\pi, \alpha, \tau)}{\partial \tau} = \frac{-(1-\pi)\pi((1-\alpha) + \delta V + (1-\alpha)\delta V)}{((1-\alpha)(1-\pi) + \pi + \tau(\delta V - \pi(1 + 2\delta V)))^2} < 0,$$

then $\kappa < \underline{\kappa}(\pi, \alpha, \tau') < \underline{\kappa}(\pi, \alpha, 0)$. □

Proof of Proposition 3. Let $\kappa, \alpha, \pi, \tau$ induce informed populism:

$$\underline{\eta}(\pi) < \eta < \bar{\eta}(\pi, \kappa) \text{ and } \kappa < \underline{\kappa}(\pi, \alpha, \tau). \tag{B.18}$$

Whenever (B.18) holds, in each period, the voter drains the swamp with probability

$$\tau \Pr[s_t = 0] = \tau(\pi(1-\kappa) + (1-\pi)\kappa). \tag{B.19}$$

Marginal changes of α such that (B.18) continues to hold will have no effect on the probability (B.19). Taking the derivatives of (B.19) with respect to κ and τ yields $\tau(1-2\pi) > 0$ and $\pi(1-\kappa) + (1-\pi)\kappa > 0$, respectively. Thus, marginal decreases in κ or τ such that (B.18) continues to hold decrease (B.19). □

Proof of Proposition 4. Let $\kappa, \alpha, \pi, \tau$ induce preemptive populism:

$$\eta < \underline{\eta}(\pi) \text{ and } \kappa < \underline{\kappa}(\pi, \alpha, \tau). \tag{B.20}$$

Whenever (B.20) holds, in each period, the voter drains the swamp with probability

$$\Pr[s_t = 0]p^*(0)(\tau + (1-\tau)d^*(\emptyset)) + \Pr[s_t = 1](1-\tau)d^*(\emptyset) = \Pr[s_t = 0]p^*(0)\tau + (1-\tau)d^*(\emptyset) \tag{B.21}$$

where, by (B.10),

$$d^*(\emptyset) = 1 - \frac{(1-\alpha)(1-\mu_t^*(0)) + \delta V}{(1-\tau)(\mu_t^*(0) + \delta V)} \in (0, 1)$$

and, by (B.9),

$$p^*(0) = \frac{\pi\kappa - \frac{1-\eta}{2}\Pr[s_t = 1]}{\frac{1-\eta}{2}\Pr[s_t = 0] - \pi(1-\kappa)} \in (0,1).$$

Substituting $d^*(\emptyset)$ into (B.21) gives that the probability that in each period the voter drains the swamp is

$$\tau \Pr[s_t = 0]p^*(0) + (1-\tau) - \frac{(1-\alpha)(1-\mu^*(0)) + \delta V}{(\mu^*(0) + \delta V)}. \tag{B.22}$$

The derivative of (B.22) with respect to τ and α are $\Pr[s_t = 0]p^*(0) - 1 < 0$ and $\frac{(1-\mu_t^*(0))+\delta V}{(\mu_t^*(0)+\delta V)} > 0$, respectively. Thus, marginal increases in τ or marginal decreases in α, such that (B.20) continues to hold, decrease (B.22).

Now we consider the effect of marginal changes in κ. Taking the derivative of (B.21) gives

$$\tau \frac{\partial \Pr[s_t = 0]p^*(0)}{\partial \kappa} + (1-\tau)\frac{\partial d^*(\emptyset)}{\partial \kappa} = \tau\Big(p^*(0)\frac{\partial \Pr[s_t = 0]}{\partial \kappa} + \Pr[s_t = 0]\frac{\partial p^*(0)}{\partial \kappa}\Big) + (1-\tau)\frac{\partial d^*(\emptyset)}{\partial \kappa}, \tag{B.23}$$

where $\frac{\partial \Pr[s_t=0]}{\partial \kappa} = 1 - 2\pi > 0$,

$$\frac{\partial d^*(\emptyset)}{\partial \kappa} = \frac{(1-\alpha+(2-\alpha)\delta V)}{(1-\tau)(\delta V + \mu_t^*(0))^2}\frac{\partial \mu_t^*(0)}{\partial \kappa} < 0,$$

and, because $\eta < \underline{\eta}(\pi) = 1 - 2\pi$,

$$\frac{\partial p^*(0)}{\partial \kappa} = \frac{(1-2\pi-\eta)(1-\eta(1-2\pi))}{(\kappa-(1+\eta)\pi-\kappa\eta(1-2\pi))^2} > 0.$$

While maintaining conditions (B.20), the derivative (B.23) may be either increasing, decreasing, or non-monotonic in κ. We illustrate these comparative statics with three specific parameter values—via a continuity argument, it follows that these three comparative statics hold for a neighborhood of parameter values. First, when $(\alpha, \pi, \delta, \tau, V, \eta) = (0.85, 0.45, 0.25, 0.7, 0.05, 0.025)$, Condition (B.20) becomes $\eta = 0.025 < \underline{\eta} = 0.1$, and $\kappa < \underline{\kappa} \approx 0.61863$, and (B.23) is positive for all $\kappa \in (1-\pi, \underline{\kappa})$. Second, when $(\alpha, \pi, \delta, \tau, V, \eta) = (0.95, 0.25, 0.25, 0.01, 0.25, 0.05)$, Condition (B.20) becomes $\eta = 0.05 < \underline{\eta} = 0.5$ and $\kappa < \underline{\kappa} \approx 0.868121$, and (B.23) is negative for all $\kappa \in (1-\pi, \underline{\kappa})$. Third, when $(\alpha, \pi, \delta, \tau, V, \eta) = (0.95, 0.4, 0.25, 0.1, 0.2, 0.05)$, Condition (B.20) becomes $\eta = 0.05 < \underline{\eta} = 0.2$ and $\kappa < \underline{\kappa} \approx$

0.915601, and (B.23) is humped-shaped for $\kappa \in (1-\pi, \underline{\kappa})$. $\square$

Proof of Proposition 5. From (B.19), it is immediate that, in informed populism, marginal changes to η have no effect on the frequency that the voter drains the swamp. In preemptive populism, in each period, the voter drains the swamp with probability (B.21). The derivative of (B.21) with respect to η is $\tau \Pr[s_t = 0] \frac{\partial p^*}{\partial \eta}$, where

$$\frac{\partial p^*}{\partial \eta} = \frac{2(2\kappa - 1)(1-\pi)\pi}{(\kappa - (1+\eta)\pi - \kappa\eta(2\pi - 1))^2} > 0.$$

Thus, marginal increases in η such that (B.20) continues to hold increase the probability (B.21). $\square$

C Additional proofs

Proposition C.1. *Let $\kappa > \bar{\kappa}(\alpha, \pi)$. In the unique equilibrium, $p^*(s) = s$ and, on the equilibrium path, the voter chooses not to drain the swamp.*

Proof. Let $\kappa > \bar{\kappa}(\alpha, \pi)$. It is straightforward to see that there exists an equilibrium whereby the agent chooses $p_t = 1$ if and only if $s_t = 1$. In this equilibrium, the voter's strategy is such that: (i) if $\eta < \bar{\eta}(\pi, \kappa)$, the voter chooses $d_t = 1$ if and only if she is informed that $s_t = 0$; (ii) if $\eta > \bar{\eta}(\pi, \kappa)$, the voter always chooses $d_t = 0$. We prove that in each case this is the unique equilibrium.

First notice that Lemmas 2 and 3 hold verbatim but Lemma 4 does not—instead, we have Lemma C.1 below.

Lemma C.1. *Let $\kappa > \bar{\kappa}(\alpha, \pi)$. In any equilibrium, $p^*(s) = s$.*

Proof of Lemma C.1. First, let $\eta > \bar{\eta}(\pi, \kappa)$. By Lemma 3, for any s_t, the voter chooses $d_t = 0$ when informed. When uninformed, the voter's belief is contained in the interval $[\nu_t(0), \nu_t(1)]$ (see (A.5)) and, hence, we also have $d_t = 0$. That is, the voter never drains the swamp. Therefore, the agent optimally chooses $p_t = 1$ if and only if $s_t = 1$ (Lemma 1).

Second, let $\eta < \bar{\eta}(\pi, \kappa)$. By Lemma 3, in every equilibrium $d^*(0) = 1$ and $d^*(1) = 0$. The auxiliary lemmas in the proof of Lemma 4, Lemmas B.1–B.4, hold verbatim in the current setting. Therefore, one of three must hold: (i) $p^*(0) = 0$ and $p^*(1) = 1$; (ii) $p^*(0) \in (0, 1)$ and $p^*(1) = 1$; or (iii) $p^*(0) = 1$ and $p^*(1) = 1$. But notice that, when $s_t = 0$, choosing $p_t = 1$ is a strictly dominated action for the agent. Thus, it must be that $p^*(0) = 0$ and $p^*(1) = 1$. $\square$

We now return to the proof of Proposition C.1. Given Lemma C.1, the voter's belief when uninformed is $\nu_t(\emptyset) = \nu_t(1)$ and, hence, he chooses $d_t = 0$. Thus, on the equilibrium path, the voter never drains the swamp. $\square$

D Endogenous V

We now study the model introduced in Section 7 and in which the agent is forward looking, (potentially) infinitely lived, maintaining the assumption that, if the voter drains the swamp, the incumbent agent ceases to live and is immediately replaces by a new (novice) agent. The agent's continuation payoff if the voter does not drain the swamp is then given by $V(\sigma)$.

In this extension, the only dynamic consequence of the agent's action is via the voter's decision to drain the swamp. Once the voter drains the swamp, the incumbent agent is replaced by a new agent and, hence, obtains payoff zero in all future periods. Thus, for a given set of parameters and assessment σ, the agent's continuation payoff is a fixed value and can be treated as an exogenous parameter. Thus, Lemmas 2–4 hold verbatim in this extension and the proof arguments are identical.

For our main result (Proposition D.1), we make the following technical assumption. As per Remark D.1, this assumption is satisfied whenever δ is sufficiently small.

Assumption 3. *Let* $\tilde{f}(d^*) = \frac{\Pr[s_t=1](\tau+(1-\tau)(1-d^*))\mu^*(1)+\Pr[s_t=0](1-\tau)(1-d^*)\mu^*(0)}{1-\delta\left(\Pr[s_t=1](\tau+(1-\tau)(1-d^*))+\Pr[s_t=0](1-\tau)(1-d^*)\right)}$. *The function*

$$\Phi(d^*) := (1-\alpha)(1-\mu(0)) + \delta\tilde{f}(d^*) - (1-\tau)(1-d^*)(\mu(0)+\delta\tilde{f}(d^*)) \tag{D.1}$$

is increasing in d^*.

Remark D.1. *A sufficient for Assumption 3 to hold is that* δ *is small; in particular,*

$$\delta < \bar{\delta}(\alpha,\tau) := \frac{1-\alpha}{2-\alpha}\frac{\tau^2}{(2+\tau)}. \tag{D.2}$$

Lemma D.1 states and proves this result.

Under Assumption 3—or alternatively (D.2)—we prove a variant of Proposition 1 below (Proposition D.1). Note that the cutoff values $\bar{\kappa}(\pi,\alpha), \underline{\eta}(\pi)$, and $\bar{\eta}(\pi,\kappa)$ in Proposition D.1 are the same as in Proposition 1.

Proposition D.1. *Suppose Assumption 3. There exists cutoffs* $\underline{\kappa}_{end}(\pi,\alpha,\tau) < \bar{\kappa}(\pi,\alpha)$ *and* $\underline{\eta}(\pi) < \bar{\eta}(\pi,\kappa)$ *such that, in equilibrium,*

(i) $\bar{\eta}(\pi,\kappa) < \eta$ *induces a technocracy;*

(ii) $\eta < \bar{\eta}(\pi,\kappa)$ *and* $\underline{\kappa}_{end}(\pi,\alpha,\tau) < \kappa$ *induces a responsive democracy;*

(iii) $\underline{\eta}(\pi) < \eta < \bar{\eta}(\pi,\kappa)$ *and* $\kappa < \underline{\kappa}_{end}(\pi,\alpha,\tau)$ *induces informed populism;*

(iv) $\eta < \underline{\eta}(\pi)$ *and* $\kappa < \underline{\kappa}_{end}(\pi,\alpha,\tau)$ *induces preemptive populism.*

Proof.

Part (i): Follows from Lemmas 2–4.

For the remainder of the proof, we suppose $\eta < \bar{\eta}(\pi,\kappa)$. By Lemma 4, the agent chooses $p_t = 1$ when $s_t = 1$ and, by Lemma 3, the voter chooses $d_t = 1$ (resp., $d_t = 0$) when informed that $s_t = 0$ (resp., $s_t = 1$).

For Parts (ii)–(iv), we characterize the set of parameters for which each regimes exists. The interior of each set of parameters will be distinct and, hence, it follows that the equilibrium is unique with exception of a measure zero set of parameters.

Part (ii): In a responsive democracy, $d^* = 0$ and $p^* = 0$. By Bayes' rule, $\nu_t^*(s_t = \emptyset) = \nu_t^*(s_t = 1)$ and, hence, the voter optimally chooses $d_t = 0$ when uninformed. The agent's strategy $p^* = 0$ is optimal if and only if

$$(1-\alpha)(1-\mu(0)) + \delta V(\sigma^*) \geq (1-\tau)(\mu(0) + \delta V(\sigma^*)), \tag{D.3}$$

where

$$V(\sigma^*) = \frac{\Pr[s_t = 1]\mu^*(1) + \Pr[s_t = 0](1-\alpha)(1-\mu^*(0))}{1-\delta} = \frac{\pi\kappa + (1-\alpha)(1-\pi)\kappa}{1-\delta}. \tag{D.4}$$

Substituting (D.4) into (D.3) and rearranging gives an equivalent condition for (D.3):

$$(1-\alpha)(1-\mu(0)) + \delta\tau\kappa\frac{\pi + (1-\alpha)(1-\pi)}{1-\delta} \geq (1-\tau)\mu(0). \tag{D.5}$$

Both sides of Inequality (D.5) are continuous in κ and the LHS (resp., RHS) is increasing (resp., decreasing) in κ. Furthermore, at $\kappa = \bar{\kappa}(\pi,\alpha)$, we have

$$(1-\alpha)(1-\mu(0)) = \mu(0),$$

and, hence, the LHS of (D.5) exceeds the RHS. Define the cutoff $\underline{\kappa}_{RD}(\pi,\alpha,\tau) \in (1-\pi, \bar{\kappa}(\pi,\alpha))$ such that (D.5) holds with equality at $\kappa = \underline{\kappa}_{RD}(\pi,\alpha,\tau)$ when such a value exists and, otherwise, define $\underline{\kappa}_{RD}(\pi, \alpha, \tau) = 1 - \pi$. Thus, for all $\kappa \in (1-\pi, \bar{\kappa}(\pi, \alpha))$, (D.5) holds if and only if $\kappa \geq \underline{\kappa}_{RD}(\pi, \alpha, \tau)$. Therefore, a responsive democracy is an equilibrium if and only if $\underline{\kappa}_{RD}(\pi, \alpha, \tau) \leq \kappa < \bar{\kappa}(\pi, \alpha)$ and $\eta < \bar{\eta}(\pi, \kappa)$.

Part (iii): In informed populism, $d^* = 0$ and $p^* = 1$. By Bayes' rule $\nu_t^*(s_t = \emptyset) = \pi$. By Lemma 2, the voter's strategy $d^* = 0$ is optimal if and only if

$$\pi \geq \frac{1-\eta}{2} \iff \eta \geq \underline{\eta}(\pi) = 1 - 2\pi. \tag{D.6}$$

Abusing notation slightly, but for greater clarity, we denote the agent's continuation payoff as $\hat{V}(\sigma^*)$. The agent's strategy $p^* = 1$ is optimal if and only if

$$(1-\alpha)(1-\mu(0)) + \delta\hat{V}(\sigma^*) \leq (1-\tau)(\mu(0) + \delta\hat{V}(\sigma^*)), \tag{D.7}$$

where

$$\hat{V}(\sigma^*) = \frac{\Pr[s_t = 1]\mu^*(1) + \Pr[s_t = 0](1-\tau)\mu^*(0)}{1 - \delta(\Pr[s_t = 1] + \Pr[s_t = 0](1-\tau))} = \frac{\pi\kappa + (1-\tau)\pi(1-\kappa)}{1 - \delta(1 - \tau(\pi(1-\kappa) + (1-\pi)\kappa))}. \tag{D.8}$$

It can be shown that the difference between the LHS and RHS of (D.7) is increasing in κ. To see this, substitute (D.8) into (D.7) and rearrange to obtain the difference between the LHS and RHS of (D.7):

$$\Psi(\kappa) := (1-\alpha) - (2-\alpha-\tau)\frac{\pi(1-\kappa)}{\pi(1-\kappa) + (1-\pi)\kappa} + \tau\delta\frac{\pi\kappa + (1-\tau)\pi(1-\kappa)}{1 - \delta(1 - \tau(\pi(1-\kappa) + (1-\pi)\kappa))}.$$

Taking the derivative gives

$$\frac{\partial\Psi(\kappa)}{\partial\kappa} = \frac{(1-\alpha)(1-\pi)\pi}{(\kappa(1-2\pi)+\pi)^2} + \frac{(1-\pi)\pi(1-\tau)}{(\kappa(1-2\pi)+\pi)^2} + \frac{\delta\pi(1-\delta(1-\pi)(2-\tau))\tau^2}{(1-\delta+\delta(\kappa(1-2\pi)+\pi)\tau)^2}. \tag{D.9}$$

We prove that $\frac{\partial\Psi(\kappa)}{\partial\kappa} > 0$. For sake of contradiction, suppose (D.9) is non-positive. Since the first 2 terms of (D.9) are positive,

$$\frac{\delta\pi(1-\delta(1-\pi)(2-\tau))\tau^2}{(1-\delta+\delta(\kappa(1-2\pi)+\pi)\tau)^2} < 0. \tag{D.10}$$

Because (D.10) holds, $(1-\delta(1-\pi)(2-\tau))<0$ and $(1-\pi)(2-\tau)>0$ and, hence, (D.10) is bounded above (D.10) when evaluated at $\delta=1$. Thus, it must be that (D.9) is also negative at $\delta=1$—substituting $\delta=1$ into gives

$$\begin{aligned}\frac{(1-\alpha)(1-\pi)\pi}{(\kappa(1-2\pi)+\pi)^2}+\frac{(1-\pi)\pi(1-\tau)}{(\kappa(1-2\pi)+\pi)^2}+\frac{\pi(1-(1-\pi)(2-\tau))}{(\kappa(1-2\pi)+\pi)^2}\\ =\frac{(1-\alpha)(1-\pi)\pi}{(\kappa(1-2\pi)+\pi)^2}+\frac{\pi^2}{(\kappa(1-2\pi)+\pi)^2},\end{aligned}$$

but this is clearly positive—a contradiction.

Combining the fact that, at $\kappa=\bar{\kappa}(\pi,\alpha)$, the LHS of (D.7) exceeds the RHS and $\frac{\partial\Phi}{\partial\kappa}>0$, allows for the following. Define the cutoff $\underline{\kappa}_{Pop}(\pi,\alpha,\tau)\in(1-\pi,\bar{\kappa}(\pi,\alpha))$ such that (D.7) holds with equality at $\kappa=\underline{\kappa}_{Pop}(\pi,\alpha,\tau)$ when such a value exists and, otherwise, define $\underline{\kappa}_{Pop}(\pi,\alpha,\tau)=1-\pi$. Then, for all $\kappa\in(1-\pi,\bar{\kappa}(\pi,\alpha))$, (D.7) holds if and only if $\kappa\le\underline{\kappa}_{Pop}(\pi,\alpha,\tau)$. Therefore, a informed populism is an equilibrium if and only if $\kappa\le\underline{\kappa}_{Pop}(\pi,\alpha,\tau)<\bar{\kappa}(\pi,\alpha)$ and $\underline{\eta}(\pi)\le\eta<\bar{\eta}(\pi,\kappa)$.

Part (iv): In preemptive populism, $d^*\in(0,1)$ and $p^*\in(0,1)$. By Bayes' rule,

$$\nu_t^*(s_t=\emptyset)=\frac{\pi\kappa+\pi(1-\kappa)p^*}{\Pr[s_t=1]+\Pr[s_t=0]p^*}.$$

and , hence, by Lemma 2, the voter's strategy $d^*\in(0,1)$ is optimal if and only if $p^*\in(0,1)$ satisfies

$$\frac{\pi\kappa+\pi(1-\kappa)p^*}{\Pr[s_t=1]+\Pr[s_t=0]p^*}=\frac{1-\eta}{2}. \tag{D.11}$$

The LHS of (D.11) is continuous and decreasing in p^* with range $(\pi,\Pr[\theta_t=1\mid s_t=1])$, and the LHS is strictly greater than the RHS at $p^*=0$. Thus, a (unique) solution $p^*\in(0,1)$ to (D.11) exists if and only if the LHS is strictly less than the RHS at $p^*=1$:

$$\eta<\underline{\eta}(\pi).$$

Abusing notation slightly, but for greater clarity, we denote the agent's continuation payoff as $\tilde{V}(d^*,\sigma^*)$. The agent's strategy $p^*\in(0,1)$ is optimal if and only if $d^*\in(0,1)$ satisfies

$$(1-\alpha)(1-\mu(0))+\delta\tilde{V}(d^*,\sigma^*)=(1-\tau)(1-d^*)(\mu(0)+\delta\tilde{V}(d^*,\sigma^*)), \tag{D.12}$$

where

$$\tilde{V}(d^*,\sigma^*) = \frac{\Pr[s_t=1](\tau+(1-\tau)(1-d^*))\mu^*(1)+\Pr[s_t=0](1-\tau)(1-d^*)\mu^*(0)}{1-\delta\Big(\Pr[s_t=1](\tau+(1-\tau)(1-d^*))+\Pr[s_t=0](1-\tau)(1-d^*)\Big)}. \tag{D.13}$$

By Assumption 3 and noting that $\tilde{f}(d^*) = \tilde{V}(d^*,\sigma^*)$, the difference between the LHS and RHS of (D.12) is increasing in d^*. Since the LHS exceeds the RHS at $d^*=1$, Equation (D.12) has a unique solution in $d^* \in (0,1)$ if and only if at $d^*=0$ the LHS is strictly less than the RHS:

$$(1-\alpha)(1-\mu(0)) + \delta\tilde{V}(0,\sigma^*) < (1-\tau)(\mu(0)+\delta\tilde{V}(0,\sigma^*)). \tag{D.14}$$

But notice that

$$\tilde{V}(0,\sigma^*) = \hat{V}(\sigma^*), \tag{D.15}$$

where $\hat{V}(\sigma^*)$ is defined as per (D.8). Thus, (D.14) simplifies to

$$(1-\alpha)(1-\mu(0)) + \delta\hat{V}(\sigma^*) < (1-\tau)(\mu(0)+\delta\hat{V}(\sigma^*)),$$

i.e., Inequality (D.7) holds strictly. As implied by the arguments of the proof of Part (iii), for $\kappa \in (1-\pi, \bar{\kappa}(\pi,\alpha))$, the cutoff $\underline{\kappa}_{Pop}(\pi,\alpha,\tau)$ is such that the inequality (D.7) holds strictly if and only if $\kappa < \underline{\kappa}_{Pop}(\pi,\alpha,\tau)$. Therefore, preemptive populism is an equilibrium if and only if $\eta < \underline{\eta}(\pi)$ and $\kappa < \underline{\kappa}_{Pop}(\pi,\alpha,\tau)$.

Final step. To complete the proof, we show $\underline{\kappa}_{RD}(\pi,\alpha,\tau) = \underline{\kappa}_{Pop}(\pi,\alpha,\tau)$. First we prove

$$\underline{\kappa}_{RD}(\pi,\alpha,\tau) \leq \underline{\kappa}_{Pop}(\pi,\alpha,\tau).$$

Let $\kappa = \underline{\kappa}_{Pop}(\pi,\alpha,\tau)$. We consider 2 cases. First, suppose that $1-\pi < \underline{\kappa}_{Pop}(\pi,\alpha,\tau)$; this means that

$$(1-\alpha)(1-\mu(0)) + \delta\hat{V}(\sigma^*) = (1-\tau)(\mu(0)+\delta\hat{V}(\sigma^*)).$$

Second, suppose that $1-\pi = \underline{\kappa}_{Pop}(\pi,\alpha,\tau)$; this means that

$$(1-\alpha)(1-\mu(0)) + \delta\hat{V}(\sigma^*) \geq (1-\tau)(\mu(0)+\delta\hat{V}(\sigma^*)).$$

By Assumption 3 and recalling that $\tilde{f}(0) = \tilde{V}(0, \sigma^*) = \hat{V}(\sigma^*)$ (see (D.15)), either case above violates (D.14) and, hence, there is no solution $d^* \in (0,1)$ to (D.12). Furthermore, since the LHS of (D.12) is strictly greater than the RHS at $d^* = 1$, it must be that, at $d^* = 0$, we have

$$(1-\alpha)(1-\mu(0)) + \delta\tilde{V}(0,\sigma^*) \geq (1-\tau)(\mu(0) + \delta\tilde{V}(0,\sigma^*)) \tag{D.16}$$

The continuation payoff $\tilde{V}(d^*, \sigma^*)$ is defined in (D.13) but, using the equilibrium indifference condition, the continuation payoff has an equivalent form:

$$\tilde{V}_{alt}(d^*,\sigma^*) = \frac{\Pr[s_t=1](\tau+(1-\tau)(1-d^*))\mu^*(1) + \Pr[s_t=0](1-\alpha)(1-\mu^*(0))}{1-\delta\Big(\Pr[s_t=1](\tau+(1-\tau)(1-d^*)) + \Pr[s_t=0]\Big)}. \tag{D.17}$$

At $d^* = 0$, we have

$$\tilde{V}(0,\sigma^*) = \tilde{V}_{alt}(0,\sigma^*) = V(\sigma^*), \tag{D.18}$$

where $V(\sigma^*)$ is the continuation payoff of the agent in the responsive democracy equilibrium. Thus, substituting (D.18) into (D.16), we have

$$(1-\alpha)(1-\mu(0)) + \delta V(\sigma^*) \geq (1-\tau)(\mu(0) + \delta V(\sigma^*)),$$

i.e., (D.5) holds. This, by construction of $\underline{\kappa}_{RD}(\pi,\alpha,\tau)$, implies that $\kappa \geq \underline{\kappa}_{RD}(\pi,\alpha,\tau)$; that is, $\underline{\kappa}_{Pop}(\pi,\alpha,\tau) \geq \underline{\kappa}_{RD}(\pi,\alpha,\tau)$.

We now prove $\underline{\kappa}_{Pop}(\pi,\alpha,\tau) \leq \underline{\kappa}_{RD}(\pi,\alpha,\tau)$. For sake of contradiction, suppose $\underline{\kappa}_{Pop}(\pi,\alpha,\tau) > \underline{\kappa}_{RD}(\pi,\alpha,\tau)$. Let $\kappa \in (\underline{\kappa}_{RD}(\pi,\alpha,\tau), \underline{\kappa}_{Pop}(\pi,\alpha,\tau))$, which implies that

$$(1-\alpha)(1-\mu(0)) + \delta V(\sigma^*) \geq (1-\tau)(\mu(0) + \delta V(\sigma^*)) \tag{D.19}$$

and

$$(1-\alpha)(1-\mu(0)) + \delta\hat{V}(\sigma^*) < (1-\tau)(\mu(0) + \delta\hat{V}(\sigma^*)). \tag{D.20}$$

Recalling (D.15) and (D.18), Inequality (D.20) becomes:

$$(1-\alpha)(1-\mu(0)) + \delta V(\sigma^*)) < (1-\tau)(\mu(0) + \delta V(\sigma^*)),$$

which contradicts (D.19). Defining $\underline{\kappa}_{end}(\pi,\alpha,\tau) = \underline{\kappa}_{RD}(\pi,\alpha,\tau) = \underline{\kappa}_{Pop}(\pi,\alpha,\tau)$ completes the proof of the proposition. □

Lastly, we prove the claim made in Remark D.1.

Lemma D.1. *Suppose* $\delta < \bar{\delta}(\alpha, \tau) := \frac{1-\alpha}{2-\alpha} \frac{\tau^2}{(2+\tau)}$. *Equation (D.1) is increasing in* d^*.

Proof. Suppose $\delta < \bar{\delta}(\alpha, \tau)$. We prove that $\frac{\partial \Phi(d^*)}{\partial d^*} > 0$ by taking a sequence of inequalities that establish a positive lower bound on the derivative. First note that

$$\begin{aligned}
\frac{\partial \Phi(d^*)}{\partial d^*} &= \delta \frac{\partial \tilde{f}(d^*)}{\partial d^*} + (1-\tau)\mu(0) + (1-\tau)\delta \tilde{f}(d^*) - (1-\tau)(1-d^*)\delta \frac{\partial \tilde{f}(d^*)}{\partial d^*} \\
&= (1-\tau)\mu(0) + (1-\tau)\delta \tilde{f}(d^*) + \big(1-(1-\tau)(1-d^*)\big)\delta \frac{\partial \tilde{f}(d^*)}{\partial d^*} \\
&> (1-\tau)\mu(0) + \big(1-(1-\tau)(1-d^*)\big)\delta \frac{\partial \tilde{f}(d^*)}{\partial d^*},
\end{aligned} \tag{D.21}$$

since $\tilde{f}(d^*) > 0$. Furthermore, we have

$$\frac{\partial \tilde{f}(d^*)}{\partial d^*} = -\frac{\pi(1-\tau)(1+\delta(2\kappa-1)(1-\pi)\tau)}{(1-\delta+\delta d^* - \delta(d^*-\pi-\kappa(1-2\pi))\tau)^2} < 0; \tag{D.22}$$

thus, the lower bound (D.21) can be lowered to:

$$\frac{\partial \Phi(d^*)}{\partial d^*} > (1-\tau)\mu(0) + \delta \frac{\partial \tilde{f}(d^*)}{\partial d^*}. \tag{D.23}$$

Furthermore, we have

$$\frac{\partial^2 \tilde{f}(d^*)}{\partial (d^*)^2} = \frac{2\delta\pi(1-\tau)^2(1+\delta(2\kappa-1)(1-\pi)\tau)}{(1-\delta+\delta d^* - \delta(d^*-\pi-\kappa(1-2\pi))\tau)^3} > 0$$

and, hence,

$$\frac{\partial \tilde{f}(d^*)}{\partial d^*} > \frac{\partial \tilde{f}(d^*)}{\partial d^*}\Big|^{d^*=0} = -\frac{\pi(1-\tau)(1+\delta(2\kappa-1)(1-\pi)\tau)}{(1-\delta(1-(\pi+\kappa(1-2\pi))\tau)^2}.$$

Noting that $\frac{\partial \tilde{f}(d^*)}{\partial d^*}\Big|^{d^*=0}$ is decreasing in δ, we obtain that

$$\frac{\partial \tilde{f}(d^*)}{\partial d^*} > \frac{\partial \tilde{f}(d^*)}{\partial d^*}\Big|^{d^*=0}\Big|^{\delta=1} = -\frac{\pi(1-\tau)(1+(2\kappa-1)(1-\pi)\tau)}{((\pi+\kappa(1-2\pi))\tau)^2}.$$

The derivative of $\frac{\partial \tilde{f}(d^*)}{\partial d^*}\Big|^{d^*=0}\Big|^{\delta=1}$ with respect to π is

$$\frac{-(1-\tau)(\kappa-\pi+2\kappa\pi+(2\kappa-1)(\kappa-\pi)\tau)}{(\kappa+\pi-2\kappa\pi)^3\tau^2}<0;$$

thus, $\frac{\partial \tilde{f}(d^*)}{\partial d^*}\Big|^{d^*=0}\Big|^{\delta=1}$ is decreasing in π and

$$\frac{\partial \tilde{f}(d^*)}{\partial d^*}>\frac{\partial \tilde{f}(d^*)}{\partial d^*}\Big|^{d^*=0}\Big|^{\delta=1}\Big|^{\pi=1/2}=-\frac{(1-\tau)(2+(2\kappa-1)\tau)}{\tau^2}.$$

Notice that $\frac{\partial \tilde{f}(d^*)}{\partial d^*}\Big|^{d^*=0}\Big|^{\delta=1}\Big|^{\pi=1/2}$ is decreasing in κ and so

$$\frac{\partial \tilde{f}(d^*)}{\partial d^*}>\frac{\partial \tilde{f}(d^*)}{\partial d^*}\Big|^{d^*=0}\Big|^{\delta=1}\Big|^{\pi=1/2}\Big|^{\kappa=1}=-\frac{(1-\tau)(2+\tau)}{\tau^2}.$$

Thus, (D.23) can be lowered to $\frac{\partial \Phi(d^*)}{\partial d^*}>(1-\tau)\mu(0)-\delta\frac{(1-\tau)(2+\tau)}{\tau^2}$. Since $\mu(0)$ is decreasing in κ and $\kappa<\bar{\kappa}(\pi,\alpha)$,

$$\mu(0)>\mu(0)\Big|^{\kappa=\bar{\kappa}(\pi,\alpha)}=\frac{1-\alpha}{2-\alpha}.$$

Thus, we have

$$\frac{\partial \Phi(d^*)}{\partial d^*}>(1-\tau)\frac{1-\alpha}{2-\alpha}-\delta\frac{(1-\tau)(2+\tau)}{\tau^2}. \tag{D.24}$$

It follows that $\frac{\partial \Phi(d^*)}{\partial d^*}>0$ for all $\delta<\frac{1-\alpha}{2-\alpha}\frac{\tau^2}{(2+\tau)}=\bar{\delta}(\alpha,\tau)$, as required. □

E Multi-period accumulation

Suppose that draining the swamp leads the voter to have an inexperienced agent for $T\geq 1$ periods. That is, after draining the swamp in period t, the voter obtains expected payoff $(1-\pi)$ in each of the next T periods. At period $t+T$, the agent is experienced. Abusing notation slightly, we define $U(\sigma^*)$ as the voter's expected continuation payoff *from an experienced agent*—which we simply refer to as the voter's continuation payoff. If the voter drains the swamp in period t, then the voter's continuation payoff from the period-t inexperienced agent is

$$1-\nu_t+\sum_{t'=1}^{T-1}\delta^{t'}(1-\pi)+\delta^T U(\sigma^*).$$

In this extension, Lemma 1 holds verbatim and the proof is unchanged. We now proceed to analyze the equilibrium behavior of the voter and agent. We begin with an auxiliary lemma that bounds the voter's continuation payoff from an experienced agent.

Lemma E.1. *In every equilibrium σ^*, $\frac{1-\pi}{1-\delta} \leq U(\sigma^*) \leq \frac{\kappa+\eta}{1-\delta}$.*

Proof. Intuitively, the voter's continuation payoff is bounded by the present discounted value of the maximum expected per-period payoff they can obtain, i.e.,

$$U(\sigma^*) \leq \frac{\Pr[s_t = 0](1-\mu(0)) + \Pr[s_t = 1]\mu(1) + \eta}{1-\delta} = \frac{(1-\pi)\kappa + \pi\kappa + \eta}{1-\delta} = \frac{\kappa+\eta}{1-\delta}.$$

For the lower bound, we begin by establishing that

$$1 - \nu_t + \eta + \delta U(\sigma^*) > 1 - \nu_t + \sum_{t'=1}^{T-1} \delta^{t'}(1-\pi) + \delta^T U(\sigma^*). \tag{E.1}$$

For sake of a contradiction, suppose that (E.1) does not hold, i.e., after rearranging

$$\iff \eta + \delta U(\sigma^*) \leq \sum_{t'=1}^{T-1} \delta^{t'}(1-\pi) + \delta^T U(\sigma^*). \tag{E.2}$$

By sequential rationality,

$$\begin{aligned} U(\sigma^*) \geq\ & \Pr[p_t = 0 \mid \sigma^*](\Pr[\theta_t = 0 \mid p_t = 0, \sigma^*] + \eta + \delta U(\sigma^*)) \\ & + \Pr[p_t = 1 \mid \sigma^*]\Big(\Pr[\theta_t = 0 \mid p_t = 1, \sigma^*] + \sum_{t'=1}^{T-1} \delta^{t'}(1-\pi) + \delta^T U(\sigma^*)\Big). \end{aligned} \tag{E.3}$$

Substituting (E.2) gives

$$\begin{aligned} U(\sigma^*) &\geq \Pr[p_t = 0 \mid \sigma^*](\Pr[\theta_t = 0 \mid p_t = 0, \sigma^*] + \eta + \delta U(\sigma^*)) \\ &\quad + \Pr[p_t = 1 \mid \sigma^*]\Big(\Pr[\theta_t = 0 \mid p_t = 1, \sigma^*] + \eta + \delta U(\sigma^*)\Big) \\ &= 1 - \pi + \eta + \delta U(\sigma^*). \end{aligned}$$

Recursively applying the above inequality, we obtain

$$\begin{aligned}
& U(\sigma^*) \geq \sum_{t'=1}^{T-1} \delta^{t'-1}(1-\pi+\eta) + \delta^{T-1}U(\sigma^*) \\
\iff\ & \delta U(\sigma^*) \geq \sum_{t'=1}^{T-1} \delta^{t'}(1-\pi+\eta) + \delta^{T}U(\sigma^*) > \sum_{t'=1}^{T-1} \delta^{t'}(1-\pi) + \delta^{T}U(\sigma^*),
\end{aligned}$$

which contradicts (E.2).

Returning to (E.3) and applying (E.1) gives

$$\begin{aligned}
U(\sigma^*) \geq\ & \Pr[p_t = 0 \mid \sigma^*](\Pr[\theta_t = 0 \mid p_t = 0, \sigma^*] + \sum_{t'=1}^{T-1} \delta^{t'}(1-\pi) + \delta^{T}U(\sigma^*)) \\
& + \Pr[p_t = 1 \mid \sigma^*]\Big(\Pr[\theta_t = 0 \mid p_t = 1, \sigma^*] + \sum_{t'=1}^{T-1} \delta^{t'}(1-\pi) + \delta^{T}U(\sigma^*)\Big) \\
=\ & 1 - \pi + \sum_{t'=1}^{T-1} \delta^{t'}(1-\pi) + \delta^{T}U(\sigma^*).
\end{aligned}$$

Applying the above inequality recursively gives $U(\sigma^*) \geq (1-\pi)/(1-\delta)$. □

Lemma E.2 establishes an analog of Lemma 2. Because of the dynamic consequences, the voter has less incentive to drain the swamp. Intuitively, the size of the 'dynamic cost' depends on the equilibrium behavior of the voter and agents, σ^*.

Lemma E.2 (Voter's optimal strategy.). *In any equilibrium, the voter drains the swamp if and only if*

$$\nu_t < \frac{1-\eta}{2} - \delta\frac{1}{2}(1-\delta^{T-1})\Big(U(\sigma^*) - \frac{1-\pi}{1-\delta}\Big),$$

where $U(\sigma^*) - \frac{1-\pi}{1-\delta} \geq 0$.

Proof. The voter's expected payoff from $d_t = 1$ is

$$(1-\nu_t) + \sum_{t'=1}^{T-1} \delta^{t'}(1-\pi) + \delta^{T}U(\sigma^*) \tag{E.4}$$

and their expected payoff from $d_t = 0$ is

$$\nu_t + \eta + \delta U(\sigma^*). \tag{E.5}$$

Thus, the voter chooses $d_t = 1$ if and only if (E.4) exceeds (E.5):

$$\nu_t < \frac{1-\eta}{2} - \delta\frac{1}{2}\Big((1-\delta^{T-1})U(\sigma^*) - \sum_{t'=1}^{T-1}\delta^{t'-1}(1-\pi)\Big) \tag{E.6}$$
$$= \frac{1-\eta}{2} - \delta\frac{1}{2}(1-\delta^{T-1})\Big(U(\sigma^*) - \frac{(1-\pi)}{1-\delta}\Big)$$

We now prove that the bracketed term in the RHS of (E.6) is non-negative. For sake of contradiction, suppose it is negative:

$$U(\sigma^*) < \delta^{T-1}U(\sigma^*) + \sum_{t'=1}^{T-1}\delta^{t'-1}(1-\pi). \tag{E.7}$$

Recursively applying (E.7) gives

$$U(\sigma^*) < \delta^{T-1}\Big(\delta^{T-1}U(\sigma^*) + \sum_{t'=1}^{T-1}\delta^{t'-1}(1-\pi)\Big) + \sum_{t'=1}^{T-1}\delta^{t'-1}(1-\pi)$$
$$= \delta^{2(T-1)}U(\sigma^*) + \sum_{t'=1}^{2(T-1)}\delta^{t'-1}(1-\pi)$$
$$< \delta^{n(T-1)}U(\sigma^*) + \sum_{t'=1}^{n(T-1)}\delta^{t'-1}(1-\pi) \qquad \text{for all } n \in \mathbb{N}.$$

In the limit, as $n \to \infty$, we have $U(\sigma^*) \le \frac{1-\pi}{1-\delta}$. By Lemma E.1, it follows that $U(\sigma^*) = \frac{1-\pi}{1-\delta}$ but then substituting this into (E.7) delivers a contradiction:

$$\frac{1-\pi}{1-\delta} < \delta^{T-1}\frac{1-\pi}{1-\delta} + \frac{(1-\pi)(1-\delta^{T-1})}{1-\delta} = \frac{1-\pi}{1-\delta}.$$

□

Lemma E.3 establishes an analog of Lemma 3. Because of the dynamic consequences, the agent has more power: $\bar{\eta}_{mp}(\sigma^*) \le \bar{\eta}(\pi, \kappa)$.

Lemma E.3 (The optimal choice of an informed voter.)**.** *In any equilibrium, if the voter observes the agent's signal s_t, then*

(i) *when $s_t = 1$, the voter never drains the swamp;*

(ii) *when $s_t = 0$, there exists*

$$\bar{\eta}_{mp}(\sigma^*) := 1 - 2\frac{\pi(1-\kappa)}{\pi(1-\kappa)+(1-\pi)\kappa} - \delta(1-\delta^{T-1})\Big(U(\sigma^*) - \frac{1-\pi}{1-\delta}\Big).$$

such that the voter drains the swamp if $\eta < \bar{\eta}_{mp}(\sigma^*)$ and does not drain the swamp if $\eta > \bar{\eta}_{mp}(\sigma^*)$.

Notice that $\bar{\eta}_{mp}(\sigma^*) \leq \bar{\eta}(\pi,\kappa)$, where $\bar{\eta}(\pi,\kappa)$ is defined as per the benchmark model.

Proof. Suppose the voter is informed that $s_t = 1$. Then, because $\kappa > 1-\pi$, their belief is $\nu_t^*(1) > 1/2$ and hence, by Lemma E.2, the voter chooses $d_t = 0$. Suppose the voter is informed that $s_t = 0$. Then their belief is $\nu_t^*(0) = \Pr[\theta_t = 1 \mid s_t = 0]$. By Lemma E.2, the voter chooses $d_t = 1$ if

$$\frac{\pi(1-\kappa)}{\pi(1-\kappa)+(1-\pi)\kappa} < \frac{1-\eta}{2} - \delta\frac{1}{2}(1-\delta^{T-1})\Big(U(\sigma^*) - \frac{1-\pi}{1-\delta}\Big).$$

Rearranging the above inequality completes the proof. □

Lemma E.4 establishes an analog of Lemma 4.

Lemma E.4 (Effectiveness begets power.). *Let $\mu(s_t)$ be the agent's belief that an emergency has occurred when she observes signal s_t. In any equilibrium, if $\eta > \bar{\eta}_{mp}(\sigma^*)$, the agent always triggers the emergency policy. If $\eta < \bar{\eta}_{mp}(\sigma^*)$, the agent triggers the emergency policy with certainty when she detects a threat and otherwise with strictly positive probability only if*

$$(1-\alpha)(1-\mu(0)) + \delta V \leq (1-\tau)(1-d^*)(\mu(0)+\delta V) \tag{E.8}$$

where d^* is the equilibrium probability that the voter preemptively drains the swamp when she is not informed.

Proof. After replacing $\bar{\eta}(\pi,\kappa)$ with $\bar{\eta}_{mp}(\sigma^*)$, the proof argument is identical to Lemma 4 and, hence, omitted. □

Before proving our main characterization, we establish some preliminary lemmas that hold for any η, including $\eta = \bar{\eta}_{mp}(\sigma^*)$.

Lemma E.5. *In any equilibrium, when $s_t = 1$, the agent chooses $p_t = 1$.*

Proof. We begin by introducing some notation. Given an equilibrium and for $s \in \{0, 1, \emptyset\}$, let $d(s)^*$ be the voter's probability of chooses $d_t = 1$ when informed that $s_t = s$. By Part (i) of Lemma E.3, for any η, $d(1)^* = 0$ and, hence, $d(1)^* \leq d(0)^*$. It can be shown that $d(1)^* \leq d(0)^*$ implies that the agent chooses $p_t = 1$ when $s_t = 1$. The proof follows a similar argument as in Lemmas B.1–B.4; however, the argument must be adapted for the possibility that the voter may drain the swamp with non-unit probability $d(0)^* < 1$ when informed that $s_t = 0$. For brevity's sake, we omit these details. □

Lemma E.6. *In every equilibrium σ^*, $\frac{\pi+\eta}{1-\delta} \leq U(\sigma^*)$.*

Proof. Recall that p^* is the probability that the agent chooses $p_t = 1$ when $s_t = 0$. If $p^* = 1$, then, by sequential rationality and Lemma E.5, it is immediate that

$$U(\sigma^*) \geq \pi + \eta + \delta U(\sigma^*) \iff U(\sigma^*) \geq \frac{\pi + \eta}{1 - \delta}.$$

Otherwise, suppose $p^* \in [0, 1)$. Then, by sequential rationality and Lemma E.5,

$$\begin{aligned} U(\sigma^*) &\geq \Pr[p_t = 0 \mid \sigma^*](\Pr[\theta_t = 0 \mid p_t = 0, \sigma^*] + \eta + \delta U(\sigma^*)) \\ &\qquad + \Pr[p_t = 1 \mid \sigma^*]\Big(\Pr[\theta_t = 1 \mid p_t = 1, \sigma^*] + \eta + \delta U(\sigma^*)\Big) \\ &= \eta + \delta U(\sigma^*) + (1 - p^*)(1 - \pi)\kappa + \pi\kappa + \pi(1 - \kappa)p^* \\ &\geq \eta + \delta U(\sigma^*) + (1 - p^*)\pi(1 - \kappa) + \pi\kappa + \pi(1 - \kappa)p^* \\ &= \eta + \delta U(\sigma^*) + \pi \\ \iff U(\sigma^*) &\geq \frac{\pi + \eta}{1 - \delta}, \end{aligned}$$

where the second inequality follows because $(1 - \pi)\kappa > \pi(1 - \kappa)$. □

For our main result (Proposition E.1), we make the following technical assumption. As per Remark E.1, this assumption is satisfied whenever δ is sufficiently small.

Assumption 4. *Let $g(p^*) = \frac{\eta+\pi p^*+\kappa(1-p^*)+\Pr[s_t=0]\tau p^*((1-\mu(0))+\sum_{t'=1}^{T-1}\delta^{t'}(1-\pi)-\mu(0)-\eta)}{1-\delta+\Pr[s_t=0]\tau p^*(1-\delta^{T-1})\delta}$. The function*

$$\zeta(p^*) := \frac{\pi\kappa + \pi(1 - \kappa)p^*}{\Pr[s_t = 1] + \Pr[s_t = 0]p^*} - \frac{1 - \eta}{2} + \delta\frac{1}{2}(1 - \delta^{T-1})\Big(g(p^*) - \frac{1 - \pi}{1 - \delta}\Big) \tag{E.9}$$

is decreasing in p^*.

Remark E.1. *A sufficient for Assumption 4 to hold is that δ is small. In particular, define $\tilde{\delta}(\pi, T)$ as the unique value in $(0,1)$ such that $\tilde{\delta}(\pi, T) = \frac{2\pi(1-\pi)(1-2\pi)(1-\tilde{\delta}(\pi,T))^2}{4+(T-1)(1-\pi)}$; such a value always exists. Then Assumption 4 holds for all $\delta < \tilde{\delta}(\pi, T)$. Lemma E.8 states and proves this result.*

Under Assumption 4—or alternatively $\delta < \tilde{\delta}(\pi, T)$—we prove a variant of Proposition 1 below (Proposition E.1). Note that the cutoff values $\bar{\kappa}(\pi, \alpha)$ and $\underline{\kappa}(\pi, \alpha, \tau)$ in Proposition E.1 are the same as in Proposition 1.

Proposition E.1 (Technocracy, democracy, and populism.). *Suppose Assumption 4 holds. There exists cutoffs $\underline{\kappa}(\pi, \alpha, \tau) < \bar{\kappa}(\pi, \alpha)$ and $\bar{\eta}_{RD}(\pi, \kappa, T) \leq \underline{\eta}_{IP}(\pi, \kappa, T) < \bar{\eta}_{Tech}(\pi, \kappa, T)$ such that, in equilibrium,*

(i) *$\bar{\eta}_{Tech}(\pi, \kappa, T) < \eta$ induces a technocracy;*

(ii) *$\eta < \bar{\eta}_{RD}(\pi, \kappa, T)$ and $\underline{\kappa}(\pi, \alpha, \tau) < \kappa$ induces a responsive democracy;*

(iii) *$\underline{\eta}_{IP}(\pi, \kappa, T) < \eta < \bar{\eta}_{Tech}(\pi, \kappa, T)$ and $\kappa < \underline{\kappa}(\pi, \alpha, \tau)$ induces informed populism;*

(iv) *$\eta < \underline{\eta}_{IP}(\pi, \kappa, T)$ and $\kappa < \underline{\kappa}(\pi, \alpha, \tau)$ induces preemptive populism.*

Proof.

Part (i): In a technocracy, the voter's continuation payoff from an experienced agent is

$$U(\sigma^*) = \frac{\pi + \eta}{1 - \delta}$$

and, hence,

$$\bar{\eta}_{mp}(\sigma^*) := 1 - \delta \frac{1 - \delta^{T-1}}{1 - \delta} \Big(\eta - (1 - 2\pi) \Big) - 2 \frac{\pi(1 - \kappa)}{\pi(1 - \kappa) + (1 - \pi)\kappa}.$$

The voter's decision to choose $d_t = 0$ when informed that $s_t = 0$ is optimal if and only if $\bar{\eta}_{mp}(\sigma^*) \leq \eta$:

$$1 + \delta \frac{1 - \delta^{T-1}}{1 - \delta}(1 - 2\pi) - 2\frac{\pi(1 - \kappa)}{\pi(1 - \kappa) + (1 - \pi)\kappa} \leq \Big(1 + \delta \frac{1 - \delta^{T-1}}{1 - \delta}\Big)\eta = \frac{1 - \delta^T}{1 - \delta}\eta.$$

Rearranging gives the condition

$$\frac{1 - \delta}{1 - \delta^T} + \delta \frac{1 - \delta^{T-1}}{1 - \delta^T}(1 - 2\pi) - 2\frac{1 - \delta}{1 - \delta^T} \frac{\pi(1 - \kappa)}{\pi(1 - \kappa) + (1 - \pi)\kappa} := \bar{\eta}_{Tech}(\pi, \kappa, T) \leq \eta.$$

Thus, a technocracy is an equilibrium if and only if $\bar{\eta}_{Tech}(\pi, \kappa, T) \leq \eta$.

Part (ii): In a responsive democracy, the voter's continuation payoff is

$$U(\sigma^*) = \frac{\Pr[s_t = 0](1-\mu(0)) + \Pr[s_t = 1]\mu(1) + \eta}{1-\delta} = \frac{(1-\pi)\kappa + \pi\kappa + \eta}{1-\delta} = \frac{\kappa + \eta}{1-\delta}.$$

It is optimal for the voter to choose $d_t = 1$ when informed that $s_t = 0$ if and only if $\eta \leq \bar{\eta}_{mp}(\sigma^*)$; that is,

$$\begin{aligned}
&\eta \leq 1 - 2\frac{\pi(1-\kappa)}{\pi(1-\kappa)+(1-\pi)\kappa} - \delta(1-\delta^{T-1})\frac{\kappa+\eta-(1-\pi)}{1-\delta} \\
\iff \eta(1+\frac{\delta(1-\delta^{T-1})}{1-\delta}) &\leq 1 - 2\frac{\pi(1-\kappa)}{\pi(1-\kappa)+(1-\pi)\kappa} - \delta(1-\delta^{T-1})\frac{\kappa-(1-\pi)}{1-\delta} \\
\iff \eta &\leq \frac{1-\delta}{1-\delta^T} - 2\frac{1-\delta}{1-\delta^T}\frac{\pi(1-\kappa)}{\pi(1-\kappa)+(1-\pi)\kappa} - \frac{\delta(1-\delta^{T-1})(\kappa-(1-\pi))}{1-\delta^T} \\
&:= \bar{\eta}_{RD}(\pi,\kappa,T).
\end{aligned}$$

Rearranging shows that

$$\begin{aligned}
\bar{\eta}_{RD}(\pi,\kappa,T) &= \frac{1-\delta}{1-\delta^T} - 2\frac{1-\delta}{1-\delta^T}\frac{\pi(1-\kappa)}{\pi(1-\kappa)+(1-\pi)\kappa} + \frac{\delta(1-\delta^{T-1})(1-2\pi)}{1-\delta^T} - \frac{\delta(1-\delta^{T-1})(\kappa-\pi)}{1-\delta^T} \\
&= \bar{\eta}_{Tech}(\pi,\kappa,T) - \frac{\delta(1-\delta^{T-1})(\kappa-\pi)}{1-\delta^T},
\end{aligned}$$

and, since $\kappa > 1-\pi > \pi$, we have that $\bar{\eta}_{RD}(\pi,\kappa,T) < \bar{\eta}_{Tech}(\pi,\kappa,T)$.

The agent's strategy $p^* = 0$ is optimal if and only if

$$(1-\alpha)(1-\mu(0)) + \delta V \geq (1-\tau)(\mu(0)+\delta V). \tag{E.10}$$

Both sides of inequality (E.10) are continuous in κ and the LHS (resp., RHS) is increasing (resp., decreasing) in κ. Furthermore, at $\kappa = \bar{\kappa}(\pi,\alpha)$, the LHS of (E.10) exceeds the RHS because $(1-\alpha)(1-\mu(0)) = \mu(0)$. Define the cutoff $\underline{\kappa}(\pi,\alpha,\tau) \in (1-\pi, \bar{\kappa}(\pi,\alpha))$ such that (E.10) holds with equality at $\kappa = \underline{\kappa}(\pi,\alpha,\tau)$ when such a value exists and, otherwise, define $\underline{\kappa}(\pi,\alpha,\tau) = 1-\pi$. Thus, for all $\kappa \in (1-\pi, \bar{\kappa}(\pi,\alpha))$, (E.10) holds if and only if $\kappa \geq \underline{\kappa}(\pi,\alpha,\tau)$.

Finally, notice that the requirement for $d^* = 0$ to be optimal for the voter does not imply any additional constraints once $\kappa \geq \underline{\kappa}(\pi,\alpha,\tau)$ is imposed since this implies $p^* = 0$ and, in turn, $d^* = 0$ is optimal for the voter. Therefore, a responsive democracy is an equilibrium if and only if

$$\underline{\kappa}(\pi,\alpha,\tau) \le \kappa < \bar{\kappa}(\pi,\alpha) \text{ and } \eta \le \bar{\eta}_{RD}(\pi,\kappa,T) \tag{E.11}$$

Part (iii): In informed populism, the voter's continuation payoff is

$$\begin{aligned}
U(\sigma^*) &= \Pr[s_t=1](\mu(1)+\eta+\delta U(\sigma^*)) + \Pr[s_t=0](1-\tau)(\mu(0)+\eta+\delta U(\sigma^*)) \\
&\quad + \Pr[s_t=0]\tau\Big(1-\mu(0)+\frac{(1-\pi)(1-\delta^T)}{1-\delta}+\delta^T U(\sigma^*)\Big) \\
&= (1-\Pr[s_t=0]\tau)\eta + \pi\kappa + (1-\tau)\pi(1-\kappa) + \tau(1-\pi)\kappa \\
&\quad + \delta U(\sigma^*)(1-\Pr[s_t=0]\tau) + \Pr[s_t=0]\tau\delta^T U(\sigma^*) \\
&\quad + \Pr[s_t=0]\tau\Big(\frac{(1-\pi)(1-\delta^T)}{1-\delta}\Big)
\end{aligned}$$

and, hence,

$$\begin{aligned}
U(\sigma^*) &= \frac{1}{1-\Big(\delta(1-\Pr[s_t=0]\tau)+\Pr[s_t=0]\tau\delta^T\Big)} \\
&\times\Bigg((1-\Pr[s_t=0]\tau)\eta + \pi\kappa + (1-\tau)\pi(1-\kappa) + \tau(1-\pi)\kappa + \Pr[s_t=0]\tau\Big(\frac{(1-\pi)(1-\delta^T)}{1-\delta}\Big)\Bigg).
\end{aligned}$$

It is optimal for the voter to choose $d_t = 1$ when informed that $s_t = 0$ if and only if $\eta \le \bar{\eta}_{mp}(\sigma^*)$:

$$\begin{aligned}
\eta &\le 1-\delta(1-\delta^{T-1})U(\sigma^*) + \delta(1-\delta^{T-1})\frac{1-\pi}{1-\delta} - 2\frac{\pi(1-\kappa)}{\pi(1-\kappa)+(1-\pi)\kappa} \\
&= 1+\delta(1-\delta^{T-1})\frac{1-\pi}{1-\delta} - 2\frac{\pi(1-\kappa)}{\pi(1-\kappa)+(1-\pi)\kappa} \\
&\quad - \frac{\delta(1-\delta^{T-1})(1-\Pr[s_t=0]\tau)}{1-\Big(\delta(1-\Pr[s_t=0]\tau)+\Pr[s_t=0]\tau\delta^T\Big)}\eta \\
&\quad - \delta(1-\delta^{T-1})\frac{\Big(\pi\kappa+(1-\tau)\pi(1-\kappa)+\tau(1-\pi)\kappa+\Pr[s_t=0]\tau\Big(\frac{(1-\pi)(1-\delta^T)}{1-\delta}\Big)\Big)}{1-\Big(\delta(1-\Pr[s_t=0]\tau)+\Pr[s_t=0]\tau\delta^T\Big)}.
\end{aligned}$$

Simplifying gives

$$
\begin{aligned}
\eta \leq\ & \frac{1}{1+\frac{\delta(1-\delta^{T-1})(1-\Pr[s_t=0]\tau)}{1-\left(\delta(1-\Pr[s_t=0]\tau)+\Pr[s_t=0]\tau\delta^T\right)}} \\
& \times \Bigg(1+\delta(1-\delta^{T-1})\frac{1-\pi}{1-\delta}-2\frac{\pi(1-\kappa)}{\pi(1-\kappa)+(1-\pi)\kappa} \\
& -\delta(1-\delta^{T-1})\frac{\left(\pi\kappa+(1-\tau)\pi(1-\kappa)+\tau(1-\pi)\kappa+\Pr[s_t=0]\tau\left(\frac{(1-\pi)(1-\delta^T)}{1-\delta}\right)\right)}{1-\left(\delta(1-\Pr[s_t=0]\tau)+\Pr[s_t=0]\tau\delta^T\right)}\Bigg) \\
& := \bar{\eta}_{IP}(\pi,\kappa,T)
\end{aligned}
\tag{E.12}
$$

By Lemma E.2 and because $\nu_t^*(s_t=\emptyset)=\pi$, the voter's strategy $d^*=0$ is optimal if and only if

$$
\pi \geq \frac{1-\eta}{2}-\delta\frac{1}{2}(1-\delta^{T-1})\left(U(\sigma^*)-\frac{1-\pi}{1-\delta}\right). \tag{E.13}
$$

Simplifying gives

$$
\begin{aligned}
\pi \geq\ & \frac{1-\eta}{2}+\delta\frac{1}{2}(1-\delta^{T-1})\frac{1-\pi}{1-\delta} \\
& -\frac{\delta\frac{1}{2}(1-\delta^{T-1})}{1-\left(\delta(1-\Pr[s_t=0]\tau)+\Pr[s_t=0]\tau\delta^T\right)}\times\Bigg((1-\Pr[s_t=0]\tau)\eta+\pi\kappa \\
& \quad +(1-\tau)\pi(1-\kappa)+\tau(1-\pi)\kappa+\Pr[s_t=0]\tau\left(\frac{(1-\pi)(1-\delta^T)}{1-\delta}\right)\Bigg) \\
=\ & \frac{1}{2}+\delta\frac{1}{2}(1-\delta^{T-1})\frac{1-\pi}{1-\delta} \\
& -\left(\frac{1}{2}+\frac{\delta\frac{1}{2}(1-\delta^{T-1})}{1-\left(\delta(1-\Pr[s_t=0]\tau)+\Pr[s_t=0]\tau\delta^T\right)(1-\Pr[s_t=0]\tau)}\right)\eta \\
& -\frac{\delta\frac{1}{2}(1-\delta^{T-1})\left(\pi\kappa+(1-\tau)\pi(1-\kappa)+\tau(1-\pi)\kappa+\Pr[s_t=0]\tau\left(\frac{(1-\pi)(1-\delta^T)}{1-\delta}\right)\right)}{1-\left(\delta(1-\Pr[s_t=0]\tau)+\Pr[s_t=0]\tau\delta^T\right)}
\end{aligned}
$$

and, hence,

$$
\begin{aligned}
\eta \geq & \frac{1}{\left(\frac{1}{2}+\frac{\delta\frac{1}{2}(1-\delta^{T-1})}{1-\left(\delta(1-\Pr[s_t=0]\tau)+\Pr[s_t=0]\tau\delta^T\right)(1-\Pr[s_t=0]\tau)}\right)} \\
& \times\left(\frac{1}{2}-\pi+\delta\frac{1}{2}(1-\delta^{T-1})\frac{1-\pi}{1-\delta}\right. \\
& \left.-\frac{\delta\frac{1}{2}(1-\delta^{T-1})\left(\pi\kappa+(1-\tau)\pi(1-\kappa)+\tau(1-\pi)\kappa+\Pr[s_t=0]\tau\left(\frac{(1-\pi)(1-\delta^T)}{1-\delta}\right)\right)}{1-\left(\delta(1-\Pr[s_t=0]\tau)+\Pr[s_t=0]\tau\delta^T\right)}\right) \\
& := \underline{\eta}_{IP}(\pi,\kappa,T). \qquad \text{(E.14)}
\end{aligned}
$$

Notice that

$$
\underline{\eta}_{IP}(\pi,\kappa,T) < \bar{\eta}_{IP}(\pi,\kappa,T). \qquad \text{(E.15)}
$$

To see this, suppose σ^* induces informed populism and η is such that $\eta = \bar{\eta}_{IP}(\pi,\kappa,T)$:

$$
\begin{aligned}
& \eta = 1-\delta(1-\delta^{T-1})\left(U(\sigma^*)-\frac{1-\pi}{1-\delta}\right)-2\frac{\pi(1-\kappa)}{\pi(1-\kappa)+(1-\pi)\kappa} \\
\iff & \frac{\pi(1-\kappa)}{\pi(1-\kappa)+(1-\pi)\kappa} = \frac{1-\eta}{2}-\delta\frac{1}{2}(1-\delta^{T-1})\left(U(\sigma^*)-\frac{1-\pi}{1-\delta}\right). \qquad \text{(E.16)}
\end{aligned}
$$

The RHS of (E.16) is exactly the RHS of (E.13) and the LHS of (E.16) is strictly less than π. Thus, (E.13) is satisfied and, hence, $\underline{\eta}_{IP}(\pi,\kappa,T) < \eta = \bar{\eta}_{IP}(\pi,\kappa,T)$.

The agent's strategy $p^* = 1$ is optimal if and only if

$$
(1-\alpha)(1-\mu(0))+\delta V \leq (1-\tau)(\mu(0)+\delta V). \qquad \text{(E.17)}
$$

As implied by the arguments of the proof of Part (ii), the cutoff $\underline{\kappa}(\pi,\alpha,\tau)$ is such that inequality (E.17) holds only if $\kappa \leq \underline{\kappa}(\pi,\alpha,\tau)$. Thus, informed populism is an equilibrium if and only if $\underline{\eta}_{IP}(\pi,\kappa,T) \leq \eta < \bar{\eta}_{IP}(\pi,\kappa,T)$ and $\kappa \leq \underline{\kappa}(\pi,\alpha,\tau)$.

Part (iv): In preemptive populism, the voter's continuation payoff is

$$
\begin{aligned}
U(\sigma^*) &= \Pr[s_t = 0](1-p^*)[1-\mu(0)+\eta+\delta U(\sigma^*)] \\
&\quad + (\Pr[s_t = 0]p^* + \Pr[s_t = 1])(1-\tau)(\mu(\emptyset)+\eta+\delta U(\sigma^*)) \\
&\quad + \Pr[s_t = 1]\tau(\mu(1)+\eta+\delta U(\sigma^*)) + \Pr[s_t = 0]\tau p^*((1-\mu(0)) + \sum_{t'=1}^{T-1} \delta^{t'}(1-\pi) + \delta^T U(\sigma^*)) \\
&= \Pr[s_t = 0](1-p^*)[1-\mu(0)+\eta+\delta U(\sigma^*)] \\
&\quad + (\Pr[s_t = 0]p^* + \Pr[s_t = 1])(1-\tau)(\mu(\emptyset)+\eta+\delta U(\sigma^*)) \\
&\quad + \Pr[s_t = 1]\tau(\mu(1)+\eta+\delta U(\sigma^*)) + \Pr[s_t = 0]\tau p^*((1-\mu(0)) + \sum_{t'=1}^{T-1} \delta^{t'}(1-\pi) + \delta^T U(\sigma^*)) \\
&\quad + \Pr[s_t = 0]\tau(\mu(0)+\eta+\delta U(\sigma^*)) - \Pr[s_t = 0]\tau(\mu(0)+\eta+\delta U(\sigma^*)).
\end{aligned}
$$

Simplifying gives

$$
\begin{aligned}
U(\sigma^*) &= \eta + \delta U(\sigma^*) + (1-p^*)(1-\pi)\kappa + (1-\tau)(\pi\kappa + \pi(1-\kappa)p^*) \\
&\quad + \tau\pi\kappa + \tau p^*\pi(1-\kappa) + \Pr[s_t = 0]\tau p^*((1-\mu(0)) + \sum_{t'=1}^{T-1} \delta^{t'}(1-\pi) - \mu(0) - \eta) \\
&\quad - \Pr[s_t = 0]\tau p^*(1-\delta^{T-1})\delta U(\sigma^*) \\
&= \eta + \delta U(\sigma^*) + \pi p^* + \kappa(1-p^*) + \Pr[s_t = 0]\tau p^*((1-\mu(0)) + \sum_{t'=1}^{T-1} \delta^{t'}(1-\pi) - \mu(0) - \eta) \\
&\quad - \Pr[s_t = 0]\tau p^*(1-\delta^{T-1})\delta U(\sigma^*).
\end{aligned}
$$

Therefore,

$$
U(\sigma^*) = \frac{\eta + \pi p^* + \kappa(1-p^*) + \Pr[s_t = 0]\tau p^*((1-\mu(0)) + \sum_{t'=1}^{T-1} \delta^{t'}(1-\pi) - \mu(0) - \eta)}{1-\delta+\Pr[s_t = 0]\tau p^*(1-\delta^{T-1})\delta}. \tag{E.18}
$$

By Lemma E.2 and Bayes' rule, the voter's strategy $d^* \in (0,1)$ is optimal if and only if $p^* \in (0,1)$ satisfies:

$$
\frac{\pi\kappa + \pi(1-\kappa)p^*}{\Pr[s_t = 1] + \Pr[s_t = 0]p^*} = \frac{1-\eta}{2} - \delta\frac{1}{2}(1-\delta^{T-1})\Big(U(\sigma^*) - \frac{1-\pi}{1-\delta}\Big). \tag{E.19}
$$

At $p^* = 0$, the LHS of (E.19) exceeds the RHS and, furthermore, Assumption 4 ensures that the difference between the LHS and RHS of (E.19) is decreasing in p^*. Therefore, a unique solution $p^* \in (0,1)$ to (E.19) exists if and only if, at $p^* = 1$, the RHS of (E.19) exceeds the LHS:

$$\pi < \frac{1-\eta}{2} - \delta\frac{1}{2}(1-\delta^{T-1})\Big(U(\sigma^*)|^{p^*=1} - \frac{1-\pi}{1-\delta}\Big). \tag{E.20}$$

But, at $p^* = 1$, the continuation payoff of the voter under informed and preemptive populism are equal. Thus, (E.20) is the same as (E.13) with the inequality reversed and strict—hence, (E.20) holds if and only if

$$\eta < \underline{\eta}_{IP}(\pi, \kappa, T),$$

where $\underline{\eta}_{IP}(\pi, \kappa, T)$ is defined in (E.14).

The agent's strategy $p^* \in (0,1)$ is optimal if and only if $d^* \in (0,1)$ satisfies:

$$(1-\alpha)(1-\mu(0)) + \delta V = (1-\tau)(1-d^*)(\mu(0) + \delta V). \tag{E.21}$$

The RHS (resp., LHS) of (E.21) is decreasing (resp., constant) in d^*. The range of the RHS is $(0, (1-\tau)(\mu(0)+V))$. At $d^* = 1$, the LHS exceeds the RHS. Therefore, a (unique) solution $d^* \in (0,1)$ to (E.21) exists if and only if at $d^* = 0$ the RHS exceeds the LHS:

$$(1-\alpha)(1-\mu(0)) + \delta V \leq (1-\tau)(\mu(0) + \delta V). \tag{E.22}$$

As implied by the arguments of the proof of Part (ii), the cutoff $\underline{\kappa}(\pi, \alpha, \tau)$ is such that the inequality (E.22) holds only if $\kappa \leq \underline{\kappa}(\pi, \alpha, \tau)$. Therefore, when preemptive populism is sustained it must be that $\eta < \underline{\eta}_{IP}(\pi, \kappa, T)$ and $\kappa \leq \underline{\kappa}(\pi, \alpha, \tau)$—it is straightforward to show that, for $\kappa \in (1-\pi, \bar{\kappa}(\pi, \alpha))$, this condition is also sufficient to sustain preemptive populism.

Ordering the η thresholds: For every equilibrium σ^*,

$$\bar{\eta}_{RD}(\pi, \kappa, T) \leq \bar{\eta}_{mp}(\sigma^*) \leq \bar{\eta}_{Tech}(\pi, \kappa, T). \tag{E.23}$$

This follows because $\bar{\eta}_{mp}(\sigma^*)$ is decreasing in $U(\sigma^*)$ and, per Lemmas E.1 and E.6, the lower (resp., upper bound) on $U(\sigma^*)$ is obtained when σ^* is a technocracy (resp., responsive democracy).

Combined with (E.15), Inequality (E.23) implies that

$$\bar{\eta}_{RD}(\pi,\kappa,T) \leq \underline{\eta}_{IP}(\pi,\kappa,T) < \bar{\eta}_{IP}(\pi,\kappa,T) \leq \bar{\eta}_{Tech}(\pi,\kappa,T).$$

Finally, we prove that $\bar{\eta}_{IP}(\pi,\kappa,T) = \bar{\eta}_{Tech}(\pi,\kappa,T)$. To see this, suppose that σ^* induces informed populism and $\eta = \bar{\eta}_{IP}(\pi,\kappa,T)$. Then, by construction, the voter is indifferent between draining the swamp when becoming informed that $s_t = 0$. Thus, the voter's continuation payoff can simplified to

$$\begin{aligned} U(\sigma^*) = \Pr[s_t = 1](\mu(1) + \eta + \delta U(\sigma^*)) + \Pr[s_t = 0](1-\tau)(\mu(0) + \eta + \delta U(\sigma^*)) \\ + \Pr[s_t = 0]\tau\Big(\mu(0) + \eta + \delta U(\sigma^*)\Big), \end{aligned}$$

which is the same continuation obtained under technocracy. It is then immediate that $\bar{\eta}_{IP}(\pi,\kappa,T) = \bar{\eta}_{Tech}(\pi,\kappa,T)$. ☐

Lemma E.7. *Suppose $\eta = \bar{\eta}_{mp}(\sigma^*)$ and $\eta \notin \{\bar{\eta}_{RD}(\pi,\kappa,T), \bar{\eta}_{Tech}(\pi,\kappa,T)\}$. Then $d^* = 0$, $p^* > 0$, and voter chooses $d_t = 1$ with positive probability only when informed that $s_t = 0$. Furthermore, if $\underline{\kappa}(\pi,\alpha,\tau) < \kappa$, then*

(i) *the voter chooses $d_t = 1$ with positive but non-unit probability when informed that $s_t = 0$.*

(ii) $p^* \in (0,1)$.

Proof. We begin by introducing some notation. Given an equilibrium and for $s \in \{0,1,\emptyset\}$, let $d(s)^*$ be the voter's probability of choosing $d_t = 1$ when informed that $s_t = s$. For $s \in \{0,1\}$, let $p(s)^*$ be the agent's probability of chooses $p_t = 1$ when $s_t = s$. Part (i) of Lemma E.3 and Lemma E.5 imply that $d(1)^* = 0$ and $p(1)^* = 1$.

First, for sake of contradiction, suppose that $d(0)^* = 0$. By Bayes' rule and Lemma E.2 and because $p(1)^* = 1$, it must be that $d(\emptyset)^* = 0$. But then, by Assumption 2, sequential rationality of the agent implies that $p(0)^* = 1$. The voter's continuation payoff is then equivalent to the technocracy equilibrium continuation payoff. Since $\eta = \bar{\eta}_{mp}(\sigma^*)$, it must be that $\eta = \bar{\eta}_{Tech}(\pi,\kappa,T)$—a contradiction. We conclude that, for all $\eta < \bar{\eta}_{Tech}(\pi,\kappa,T)$ and when $\eta = \bar{\eta}_{mp}(\sigma^*)$, in equilibrium $d(0)^* > 0$.

Second, for sake of contradiction, suppose $p(0)^* = 0$. By Bayes' rule and Lemma E.2, it must be that $d(\emptyset)^* = 0$. The voter's continuation payoff is then equivalent to the responsive democracy equilibrium continuation payoff. Since $\eta = \bar{\eta}_{mp}(\sigma^*)$, it must be that $\eta = \bar{\eta}_{RD}(\pi,\kappa,T)$—a contradiction. We conclude that, for all $\eta > \bar{\eta}_{RD}(\pi,\kappa,T)$ and when $\eta = \bar{\eta}_{mp}(\sigma^*)$, in equilibrium $p(0)^* > 0$.

Third, notice that $\eta = \bar{\eta}_{mp}(\sigma^*)$ is equivalent to requiring that

$$\nu_t^*(0) = \frac{1-\eta}{2} - \delta\frac{1}{2}(1-\delta^{T-1})\Big(U(\sigma^*) - \frac{1-\pi}{1-\delta}\Big).$$

Because $p(0)^* > 0$ and $p(1)^* = 1$, it follows that $\nu_t^*(0) < \nu_t^*(\emptyset)$ and hence

$$\nu_t^*(\emptyset) > \frac{1-\eta}{2} - \delta\frac{1}{2}(1-\delta^{T-1})\Big(U(\sigma^*) - \frac{1-\pi}{1-\delta}\Big),$$

i.e., the voter strictly prefers to not drain the swamp when uninformed. We conclude that $d(\emptyset)^* = 0$.

For Parts (i) and (ii), we suppose $\underline{\kappa}(\pi,\alpha,\tau) < \kappa$. For sake of contradiction, suppose $d(0)^* = 1$. As shown above, $p(0)^* > 0$ and hence it must be that

$$(1-\alpha)(1-\mu(0)) + \delta V \leq (1-\tau)(\mu(0) + \delta V).$$

But, by construction, $\underline{\kappa}(\pi,\alpha,\tau) < \kappa$ implies that the above inequality does not hold (see the proof of Part (ii) of Proposition E.1)—a contradiction. For sake of contradiction, suppose $p(0)^* = 1$. Because $d(0)^* \in (0,1)$, the voter is indifferent between draining the swamp when informed $s_t = 0$. It follows that the voter's continuation payoff must then equal the continuation payoff in a technocracy equilibrium. But then $\eta = \bar{\eta}_{Tech}(\pi,\kappa,T)$, which was assumed not to be the case—a contradiction. □

Proposition E.2. *Suppose $\underline{\kappa}(\pi,\alpha,\tau) < \kappa$ and $\bar{\eta}_{RD}(\pi,\kappa,T) < \eta < \bar{\eta}_{Tech}(\pi,\kappa,T)$. There is a unique equilibrium such that $d^* = 0$, $p^* \in (0,1)$, and the voter chooses $d_t = 1$ with positive but non-unit probability only if they are informed that $s_t = 0$.*

Proof. Suppose $\underline{\kappa}(\pi,\alpha,\tau) < \kappa$ and $\bar{\eta}_{RD}(\pi,\kappa,T) < \eta < \bar{\eta}_{Tech}(\pi,\kappa,T)$. We show that there always exists an equilibrium with $d(0)^* \in (0,1)$, $d(\emptyset)^* = 0$ and $p^* \in (0,1)$. The voter's strategy $d(0)^* \in (0,1)$ is optimal if and only if $p^* \in (0,1)$ satisfies

$$\nu_t^*(0) = \frac{1-\eta}{2} - \delta\frac{1}{2}(1-\delta^{T-1})\Big(U(\sigma^*) - \frac{1-\pi}{1-\delta}\Big). \tag{E.24}$$

Voter's strategy $d(\emptyset)^* = 0$ is optimal if and only if $p^* \in (0,1)$ satisfies

$$\nu_t^*(\emptyset) \geq \frac{1-\eta}{2} - \delta\frac{1}{2}(1-\delta^{T-1})\Big(U(\sigma^*) - \frac{1-\pi}{1-\delta}\Big). \tag{E.25}$$

Because $\nu_t^*(0) < \nu_t^*(\emptyset)$ for all $p^* \in (0,1)$, Inequality (E.25) is implied by (E.24).

The agent's strategy $p^* \in (0,1)$ is optimal if and only if $d^*(0) \in (0,1)$ satisfies

$$(1-\alpha)(1-\mu(0)) + \delta V = ((1-\tau) + \tau(1-d(0)^*))(\mu(0) + \delta V). \tag{E.26}$$

At $d^*(0) = 1$ and because $\kappa > \underline{\kappa}(\pi, \alpha, \tau)$, we have

$$(1-\alpha)(1-\mu(0)) + \delta V > (1-\tau)(\mu(0) + \delta V), \tag{E.27}$$

i.e., the LHS of (E.26) exceeds the RHS. At $d^*(0) = 0$ and because $\kappa < \bar{\kappa}(\pi, \alpha)$, we have

$$(1-\alpha)(1-\mu(0)) + \delta V < \mu(0) + \delta V, \tag{E.28}$$

i.e., the RHS of (E.26) exceeds the LHS. Given (E.27) and (E.28) and because the LHS of (E.26) is independent of $d(0)^*$ and the RHS is decreasing in $d(0)^*$, there exists a unique solution $d(0)^* \in (0,1)$ satisfying (E.26).

It only remains to prove that there exists $p^* \in (0,1)$ satisfying (E.24). Recalling Lemma E.3, (E.24) is equivalent to $\eta = \bar{\eta}_{mp}(\sigma^*)$. Next we calculate the voter's continuation payoff

$$\begin{aligned}
U(\sigma^*) &= \Pr[s_t = 0](1-p^*)(\Pr[\theta_t = 0 \mid s_t = 0] + \eta + \delta U(\sigma^*)) \\
&\quad + (p^* \Pr[s_t = 0] + \Pr[s_t = 1])(1-\tau)(\Pr[\theta_t = 1 \mid s_t = \emptyset, \sigma^*] + \eta + \delta U(\sigma^*)) \\
&\quad + \Pr[s_t = 1]\tau(\Pr[\theta_t = 1 \mid s_t = 1] + \eta + \delta U(\sigma^*)) \\
&\quad + p^* \Pr[s_t = 0]\tau(\Pr[\theta_t = 1 \mid s_t = 0] + \eta + \delta U(\sigma^*)) \\
U(\sigma^*) &= \eta + \delta U(\sigma^*) + (1-p^*)(1-\pi)\kappa + (1-\tau)(\pi\kappa + \pi(1-\kappa)p^*) + \tau\pi\kappa + \tau p^* \pi(1-\kappa) \\
U(\sigma^*) &= \eta + \delta U(\sigma^*) + \pi p^* + \kappa(1-p^*) \\
\iff U(\sigma^*) &= \frac{\eta + \pi p^* + \kappa(1-p^*)}{1-\delta}.
\end{aligned}$$

Notice that $U(\sigma^*)$ is continuous and decreasing in p^*. At $p^* = 0$ (resp., $p^* = 1$), the continuation payoff equals the continuation payoff under responsive democracy (resp., technocracy). Thus, for all $\eta \ : \ \bar{\eta}_{RD}(\pi, \kappa, T) < \eta < \bar{\eta}_{Tech}(\pi, \kappa, T)$, there is a unique $p^* \in (0,1)$ such that $\eta = \bar{\eta}_{mp}(\sigma^*)$ and, hence satisfying (E.24). □

Lastly, we prove the claim made in Remark E.1.

Lemma E.8. *Suppose $\delta < \tilde{\delta}(\pi, T)$, where $\tilde{\delta}(\pi, T)$ is defined in Remark E.1. Equation (E.9) is decreasing in p^*.*

Proof. We now prove that $\zeta(p^*)$ is decreasing in p^*. Taking the derivative:

$$\frac{\partial \Phi(\sigma^*)}{\partial p^*} = \frac{\partial\left(\frac{\pi\kappa+\pi(1-\kappa)p^*}{\Pr[s_t=1]+\Pr[s_t=0]p^*}\right)}{\partial p^*} + \delta\frac{1}{2}(1-\delta^{T-1})\frac{\partial g(p^*)}{\partial p^*}. \tag{E.29}$$

The first term is negative

$$\begin{aligned}\frac{\partial\left(\frac{\pi\kappa+\pi(1-\kappa)p^*}{\Pr[s_t=1]+\Pr[s_t=0]p^*}\right)}{\partial p^*} &= \frac{\pi(\Pr[s_t=1]-\kappa)}{(\Pr[s_t=1]+\Pr[s_t=0]p^*)^2} \\ &= \frac{-\pi(1-\pi)(2\kappa-1)}{(\Pr[s_t=1]+\Pr[s_t=0]p^*)^2} \\ &< 0 \end{aligned} \tag{E.30}$$

Taking the derivative of $g(p^*)$, we obtain

$$\begin{aligned}&\frac{\partial g(p^*)}{\partial p^*} = \frac{1}{(1-\delta+\delta\Pr[s_t=0]p^*\tau(1-\delta^{T-1}))^2}\times \\ &\Big(\pi-\delta\pi-\kappa+\Pr[s_t=0]\tau+\delta^T\Pr[s_t=0]\tau(\kappa+\eta)+\Pr[s_t=0]\tau(-2\mu(0)-\eta+\sum_{t'=1}^{T-1}\delta^{t'}(1-\pi)) \\ &\qquad +\delta(\kappa-\Pr[s_t=0]\tau(1+\kappa-2\mu(0)+\sum_{t'=1}^{T-1}\delta^{t'}(1-\pi)))\Big);\end{aligned}$$

if the numerator is non-positive, then $\frac{\partial g(p^*)}{\partial p^*} \leq 0$ and, by (E.30), the RHS of (E.29) is negative and the proof is complete. Thus, we suppose that the numerator is positive and proceed to construct an upper bound on $\frac{\partial g(p^*)}{\partial p^*}$. First notice that

$$\begin{aligned}&\frac{\partial g(p^*)}{\partial p^*} \leq \frac{1}{(1-\delta)^2}\Big(\pi+\Pr[s_t=0]\tau+\delta^T\Pr[s_t=0]\tau(\kappa+\eta)+\Pr[s_t=0]\tau(-\eta+\sum_{t'=1}^{T-1}\delta^{t'}(1-\pi))+\delta\kappa\Big) \\ &\leq \frac{1}{(1-\delta)^2}\Big(\pi+\Pr[s_t=0]\tau+\Pr[s_t=0]\tau(1+\eta)+\Pr[s_t=0]\tau(-\eta+\sum_{t'=1}^{T-1}\delta^{t'}(1-\pi))+\delta\kappa\Big) \\ &= \frac{1}{(1-\delta)^2}\Big(\pi+\Pr[s_t=0]\tau+\Pr[s_t=0]\tau+\Pr[s_t=0]\tau\sum_{t'=1}^{T-1}\delta^{t'}(1-\pi)+\delta\kappa\Big) \\ &\leq \frac{1}{(1-\delta)^2}\Big(\pi+2+\tau\sum_{t'=1}^{T-1}\delta^{t'}(1-\pi)+\delta\Big) \\ &\leq \frac{1}{(1-\delta)^2}\Big(3+(T-1)(1-\pi)+\delta\Big)\end{aligned}$$

Furthermore,

$$\begin{aligned}\delta\frac{1}{2}(1-\delta^{T-1})\frac{\partial g(p^*)}{\partial p^*} &\leq \delta\frac{1}{2}(1-\delta^{T-1})\frac{1}{(1-\delta)^2}\Big(3+(T-1)(1-\pi)+\delta\Big)\\ &\leq \delta\frac{1}{2}\frac{1}{(1-\delta)^2}\Big(4+(T-1)(1-\pi)\Big).\end{aligned}$$

Returning to (E.29), we have that $\frac{\partial\zeta(p^*)}{\partial p^*}<0$ if

$$-\pi(1-\pi)(1-2\pi)+\delta\frac{1}{2}\frac{1}{(1-\delta)^2}\Big(4+(T-1)(1-\pi)\Big)<0$$

which simplifies to

$$\delta<\frac{2\pi(1-\pi)(1-2\pi)(1-\delta)^2}{\Big(4+(T-1)(1-\pi)\Big)} \tag{E.31}$$

Let $\tilde{\delta}(\pi,T)\in(0,1)$ be the value of δ that solves (E.31) with equality. There exists a unique value because at $\delta=0$ (resp., $\delta=1$) the LHS is less (resp., greater) than the RHS and the difference between the two sides is monotonic in δ. It follows that, for all $\delta<\tilde{\delta}(\pi,T)$, $\zeta(p^*)$ is decreasing in p^*, which completes the proof. □