Hindsight biased policy evaluation*

Florian Schuett† Alexander K. Wagner‡

February 2008

Abstract

In this paper we present a political-agency model in which voters exhibit a cognitive deficiency known as hindsight bias: after the uncertainty about an event is resolved, they consider the realized outcome more foreseeable than it actually was. For their reelection decision, voters evaluate the politician's ability based on the history of observed actions and outcomes. High ability is defined as an informational advantage over voters as to the welfare maximizing policy. This creates incentives for low-ability politicians to deviate from the optimal policy choice in an attempt to be perceived as possessing superior private information. We show that, because hindsight biased voters are less impressed than rational voters when a risky policy succeeds in spite of public pessimism, the bias acts as a discipline device on low-ability politicians and may thus be welfare enhancing. It also increases political turnover compared to fully rational evaluation. While hindsight bias benefits voters in terms of politicians' discipline, its effects on selection are ambiguous. These insights may be relevant to other principal-agent relationships without ex ante commitment, e.g., promotion decisions in organizations.

Keywords: political agency, hindsight bias, behavioral economics

JEL classification numbers: D72, D83, C72, D82

---

* We are grateful to Dan Ariely, Raphaël Levy, François Salanié, Paul Seabright, Johannes Spinnewijn, Jean Tirole and Nicolas Treich for helpful discussions. We also thank participants at the 2007 meeting of the French Economic Association in Lyon, the LABSI conference in Siena, the ASSET conference in Padova, and seminar participants at Toulouse School of Economics for their comments.

† Toulouse School of Economics (GREMAQ), Manufacture des Tabacs, Aile JJ Laffont, 21 allée de Brienne, 31000 Toulouse (France), Email: florian.schuett@univ-tlse1.fr

‡ Toulouse School of Economics (LERNA), Manufacture des Tabacs, Aile JJ Laffont, 21 allée de Brienne, 31000 Toulouse (France), Email: alexander.wagner@sip.univ-tlse1.fr

1 Introduction

It is well documented in the psychology literature that many human beings exhibit hindsight bias: in retrospect, they systematically overestimate the degree to which events were predictable. The bias can be expected to cause problems in various situations where a decision maker is evaluated after the outcome of his actions is known and where there is no ex ante contract mapping outcomes into performance assessments. It is widely believed to adversely affect the legal system (Rachlinski, 1998; Viscusi, 2001; Harley, 2007); for instance, in judicial assessments of liability, hindsight bias may lead courts to hold liable even defendants who took reasonable care, given the information available at the time. Rachlinski (1998) argues that courts have responded to this problem by developing various measures to mitigate the effects of hindsight bias, and have been quite successful in containing it.[1]

In many respects, voters in a democracy have a task similar to that of jurors or judges in a court of law: much like the latter have to assess whether a defendant took the right precautions, the former need to evaluate whether an incumbent chose the right policies. Given this similarity, one might expect that hindsight bias is as problematic in the political system as in the legal system.[2] A well-documented example of how the bias can influence the public perception of policy choices, and arguably the outcome of elections, is the Gulf war of 1991. Concerning President George Bush's decision to use military force against Iraq, Mueller (1994, p. 87) reports:

> In December 1990, respondents had split about 50/50 on a question asking whether they preferred sanctions or military action. But when asked after the war how they had felt before the war, those inclined to remember that they had supported military action outnumbered those recalling their support for sanctions by nearly four to one.

If voters considered the decision in retrospect as an easy call, such a belief surely was detrimental to the President's hope of being confirmed in office.[3] This suggests that hindsight biased evaluation can have a distortionary effect on elections. Why is it then that political institutions have not evolved in ways that guard against the bias, as the law has? In this paper,

[1] Rachlinski (1998) cites several rules which aim at reducing the bias, including suppression of evidence (namely, the inadmissibility of post-accident remedial measures taken by the defendant as proof of negligence), using the profession's ex ante customs as standard of "reasonableness" (as in medical malpractice cases) and the adoption of no-liability rules (such as the business judgment rule: corporate executives generally are not liable to shareholders for decisions that turn out badly).

[2] Camerer *et al.* (1989, p. 1246), for example, voice concerns that the problem caused by hindsight bias should be "especially acute in public decision making, in which principals are a diffuse group of voters and contracts are rarely explicit." Frey and Eichenberger (1991) and Gowda (1999) also mention the bias in connection with politics.

[3] We discuss this example in more detail in section 6.

we argue that the bias may be less detrimental to the political system than to the legal system – surprisingly, we show that voters may even benefit from being hindsight biased. While it is true that hindsight bias leads to less accurate judgments, it also has a more indirect, strategic effect on the decision maker's behavior. With respect to the latter effect, democratic elections differ from court trials because politicians, unlike defendants, have reelection concerns.[4] A growing literature (reviewed below) finds that, when politicians have private information about the optimal policy, these reelection concerns may lead to various forms of inefficient signaling. We show that hindsight bias on the part of voters reduces certain wasteful activities and may therefore be welfare-enhancing. In a nutshell, the argument is that a politician may sometimes choose a policy he and the public believe to be overly risky, making it look as if he were in possession of superior private information. If the gamble pays off, rational voters are "surprised" and make an upward adjustment of their beliefs about the politician's ability. Hindsight biased voters, however, think they knew all along that the policy was going to work (they are not surprised), and do not give him as much credit. Therefore, when facing hindsight biased voters, the politician is less tempted to engage in such behavior in the first place.

Starting with the work of Fischhoff (1975), psychological research on hindsight bias has firmly established its robustness. It is not limited to college students but also affects surgeons, judges and other experts; moreover, teaching people to avoid the bias has proven tremendously difficult.[5] Given its pervasiveness, it has been suggested that hindsight bias is also likely to be relevant in a number of real-world situations without ex ante commitment. This includes the examples mentioned above, i.e., courts judging the liability of an injurer and voters in a democratic election assessing the competence of a politician, but also covers situations as diverse as human resource managers choosing whether to promote an employee or an organization deciding whether to re-appoint an expert.[6]

This paper focuses on political-economy issues.[7] In our model, a politician whose ability is unknown to voters has to choose between a risky and a safe policy. While voters and

[4] Another important difference is that liability law, which is designed to give agents incentives to take the efficient level of care, addresses a fundamental conflict of interest between society and potential injurers. By contrast, politicians' and voters' interests do not necessarily diverge as strongly because social welfare probably is an argument of many politicians' utility function (if only because they are drawn from the pool of citizens).

[5] For a review of the vast literature on hindsight bias consult Hawkins and Hastie (1990). In particular, hindsight bias has been demonstrated by Detmer *et al.* (1978) in surgeons and by Anderson *et al.* (1993) in state and federal judges. On the difficulties of debiasing see, e.g., Hoch and Loewenstein (1989).

[6] In all of those cases, hindsight bias can be expected to decrease the accuracy of the evaluation. The consequences of hindsight bias are not limited to these kinds of situations. It is also said to impair learning from the past by portfolio managers (Biais and Weber, 2006) or strategic planners (Bukszar and Connolly, 1988).

[7] Nevertheless, our analysis may be applicable to some of the other problems mentioned above, as we discuss in the conclusion.

low-ability politicians obtain only an imperfect signal of which policy is preferable ex ante, high-ability politicians know the state of the world with certainty. We assume that the public signal is informative, so efficiency requires that the low type always choose the policy suggested by the signal. This setup creates incentives for a low-ability politician to inefficiently ignore publicly available information about the welfare-maximizing policy in an attempt to "look smart", i.e., to make it seem as if he had superior private information, the trademark of competent politicians. To see why, consider what happens in case of fully rational voters, noting that a high-ability politician always chooses the right policy and thus disregards the publicly observed signal. If the low-ability politician always follows the signal, rational voters infer that any politician who chooses a policy that is contrary to the signal must be of high ability; thus, choosing an unpopular policy acts as a signal of competence. We show that if the signal is not too precise and the politician cares sufficiently about reelection, the equilibrium with fully rational voters has the low-ability politician randomizing between the policy suggested by the signal and doing exactly the opposite. Of course, this randomizing behavior is detrimental to welfare because policy choices are not optimal given the available information.

We assume that voters suffering from hindsight bias distort their recollection of the signal so as to make it consistent with the realized outcome. If the signal suggested that the safe policy was optimal, but the politician successfully enacts a risky policy, then voters wrongly believe that the signal had suggested all along that the risky policy was the right choice (they think they knew it all along). Therefore, with hindsight biased voters, some of the gain in reputation that follows from an unpopular policy which then turns out to be a success is destroyed, because ex post, biased voters think that it was the obvious choice anyway. We assume that politicians are aware of voters' behavioral decision making.[8] Anticipating voters' biased belief updating, the low-ability politician chooses a suboptimal policy less often when voters are hindsight biased then when they are perfectly rational. Thus, hindsight bias on the part of voters acts as a discipline device by reducing incentives for the low-ability politician to engage in inefficient signaling.

The disciplining effect of hindsight biased policy evaluation is unambiguously beneficial for voters' first-period welfare. However, an overall welfare assessment also has to take into account the second (i.e., post-election) period. We analyze how hindsight bias affects the selection of the second-period politician and show that it operates through two channels. First, biased voters hold erroneous posteriors about the incumbent's ability, sometimes leading them to elect the wrong candidate. Second, hindsight bias may also generate offsetting benefits in

[8] This is in line with statements from political scientists who acknowledge that "politicians typically have a strong intuitive understanding of voters' heuristics and biases" (Gowda, 1999, p. 71).

terms of inferences about the politician's type – in fact, by inducing more discipline on the low type's part, the bias can make it easier for voters to distinguish low from high-ability politicians. The overall selection effect is indeterminate: depending on parameters, hindsight bias may increase or decrease voters' second-period welfare. We nevertheless derive a polar case in which we are able to isolate the first of the two effects. When signal precision is such that differences between equilibrium strategies are small, the bias is sure to be detrimental. These qualifications notwithstanding, if voters discount future payoffs at a sufficiently high rate, hindsight bias can be welfare-enhancing regardless of what happens in the second period. Finally, we find that both the low- and the high-ability politician are less likely to be reelected when voters are hindsight biased. Hence, the bias increases political turnover.

One important departure from the political economy literature is our assumption that voters obtain an informative signal about the state of the world, just like politicians do, and that the signal coincides with that of low-ability politicians. This reflects the idea that voters are exposed to a certain amount of policy relevant public information (e.g., from the media). Note that, nevertheless, politicians *on average* still have "expertise", i.e., they are more likely to have correct information concerning the underlying state of the world than voters.

Our contribution to the literature is twofold; the first is in terms of behavioral economics, the second in terms of political economy. Our result that a behavioral bias ca improve welfare is similar in spirit to Compte and Postlewaite (2004), Bénabou and Tirole (2002) and Köszegi (2006). Those papers, however, consider how a psychological bias (namely, overconfidence) affects *intrapersonal* welfare.[9] By contrast, we investigate how a bias on the part of one group of agents (voters) can affect the behavior of other agents (politicians) in a way that increases the former's welfare.[10] Moreover, we are able to use a standard welfare measure that is unaffected by which "self" of an individual one considers; nor does it involve belief consumption.

Our paper also extends the literature on political agency,[11] in particular by going beyond the standard rational-voter assumption and instead considering behavioral decision making, as suggested by Besley (2006). Moreover, our basic model is related to the recent literature

[9] In Compte and Postlewaite (2004), an agent's self-confidence affects his performance at a task. Information-processing biases such as repressing memories of bad performance can improve the individual's welfare by boosting his confidence, thus helping him do better. Bénabou and Tirole (2002) show how over-confidence can help an individual overcome time inconsistency and thus improve his well-being (at least from an ex ante ("self zero") perspective). Köszegi (2006) lets individuals consume their self-perception, so that an overly positive self-image can raise utility.

[10] Camerer *et al.* (1989) speculate on the fact that a related phenomenon which they refer to as the "curse of knowledge" may be welfare-enhancing by (partly) eliminating market inefficiencies caused by asymmetric information. Their argument is very different from ours, though.

[11] See Persson and Tabellini (2000) or Chapter 3 in Besley (2006) for a recent overview of political agency models.

on the dysfunctional effects of electoral accountability which can arise when politicians have better information than voters. Majumdar and Mukand (2004), whose basic model is similar to ours, demonstrate that reelection concerns may lead to inefficient experimentation and policy persistence. Harrington (1993) shows that, in the presence of reelection pressures, an otherwise benevolent politician may choose the policy most likely to be well received by voters, rather than the one he himself believes to be welfare-maximizing. Similarly, in Maskin and Tirole (2004) and Smart and Sturm (2006), reelection concerns not only reduce opportunism by bad politicians, but also distort good politicians' behavior. Politicians may diverge from their preferred (and socially optimal) policy, and choose a popular policy instead, in order to signal their congruence with voters, a behavior sometimes referred to as pandering. In Dewatripont and Seabright (2006), politicians signal that they care about voters through wasteful spending on public projects. In the paper most closely related to ours, Canes-Wrone *et al.* (2001) investigate the case where politicians try to signal competence rather than congruence. They also obtain a pandering result: in their model, politicians may choose a suboptimal policy simply because it is popular among voters, provided that there is a low probability that voters learn the policy outcome before the election.

Pandering contrasts with the inefficiency in our model, where politicians signal their ability by choosing an *unpopular* policy. A number of authors obtain this kind of distortion. Canes-Wrone *et al.* (2001) derive cases for which a politician may engage in something they call "fake leadership": he acts against both popular belief and his private signal, trying to be perceived as a leader.[12] Similar results are obtained by Levy (2004) for the case of decision makers with career concerns. The decision makers in her model display a behavior labeled "anti-herding", i.e., they have a tendency to take decisions contradicting the public prior.[13] In Prat (2005), the agent may disregard valuable private information in an attempt to mimic the more able type.[14] As a result, the principal may be better off not observing the agent's action.

The finding that, for a principal, less information can be better than more is a common

[12] It is interesting that, while conventional wisdom holds that pandering on the part of politicians is common, this view is far from unanimous among political scientists; see, for example, Jacobs and Shapiro (2000) whose book is provocatively titled "Politicians Don't Pander". The authors claim that politicians' responsiveness to public opinion has been low over the past decades. This view is backed by Monroe (1998) who finds that only 55 percent of policies enacted in the US between 1980 and 1993 were consistent with the opinion of a majority of voters.

[13] Low-ability politicians' behavior in our model also bears some resemblance to Allen and Gorton (1993), where bad brokers pretend to have superior private information allowing them to identify undervalued stocks, but actually just speculate, and to models where players have a strategy labeled "gambling for resurrection" (see Downs and Rocke (1994) for a political economy application), since they are choosing a policy that is a long shot.

[14] In Prat's (2005) model, one realization of the signal is better news about the agent's type than the other (Prat calls it the "smart" realization), and the bad agent may choose the action corresponding to the "smart" realization rather than follow his true signal.

theme in setups where the principal lacks commitment power. It is a feature of Crémer (1995), for example, where the principal may forego a costless monitoring technology in order to increase the agent's incentives. In our model, incentives for the low type are improved due to voters' distorted memories; it is their hindsight bias that destroys information. Although it is well known that restrictions on information acquisition can be beneficial for the principal, our contribution is to show that a psychological bias can have such an effect.

The remainder of the paper is organized as follows. Section 2 introduces the main model while section 3 then establishes the benchmark rational-evaluation equilibrium. In section 4 we define a hindsight biased information structure, determine the equilibrium under biased policy evaluation and compare it to the rational equilibrium. A discussion follows thereafter. Selection and welfare implications of hindsight biased policy evaluation are studied in section 5. Finally, section 6 concludes. All proofs are relegated to the Appendix.

2 Basic model

In the simple two-period political agency model we consider, a politician (decision maker) chooses an action and then a possibly biased electorate (evaluator) judges the decision's quality. For the sake of clarity, we start by introducing the model under rational evaluation while assumptions for hindsight biased evaluation are deferred to section 4.

Information

In each period, the state of the world ω can be either 0 or 1. Politicians and voters receive an imperfect public signal $\sigma \in \{\sigma_0, \sigma_1\}$ about the state of the world. The probability that the signal is correct (i.e., that the state of the world is the one indicated by the signal) is $\nu_0 \equiv \Pr[\omega = 0|\sigma_0] \in (0,1)$ in case of σ_0 and $\nu_1 \equiv \Pr[\omega = 1|\sigma_1] \in (0,1)$ in case of σ_1. We will refer to ν as *signal precision*. There are two possible actions $a \in \{a_0, a_1\}$ from which the decision maker can choose. Policy a_0 is riskless and a_1 is risky.[15] The policy outcome (in terms of voters' payoff) is $y \in \{0, \Delta\}$. Action a_0 always yields a payoff of 0 to society, while action a_1 costs c and delivers a payoff of Δ if $\omega = 1$, and 0 if $\omega = 0$. We assume $\Delta > c > 0$ so that a_1 yields a higher payoff than a_0 if and only if $\omega = 1$.

A politician can either be of high (H) or low (L) ability and each politician knows his own type $\theta \in \{\theta_L, \theta_H\}$. The prior probability $\lambda_I \in (0,1)$ that the incumbent is of high ability is common knowledge. High ability is defined as an informational advantage over voters as

[15] Such an asymmetry is inherent in many policy choices, such as the decision between peace and war or between maintaining the status quo and implementing a reform. From a more pragmatic point of view, it greatly simplifies the analysis.

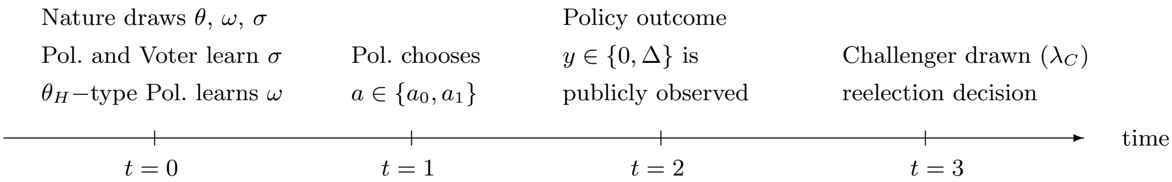

Figure 1: Timing of the game (term 1)

to the welfare-maximizing policy. While a low-ability politician only learns the signal (σ), a high-ability politician also knows the state of the world (ω) with certainty. The following assumption regarding the signal's informativeness is imposed:

Assumption 1 *Signal precision satisfies*

$$1 - \nu_0 < \frac{c}{\Delta} < \nu_1.$$

This assumption ensures that, given signal σ_ω, social welfare is maximized if the low-ability politician implements policy a_ω.

Timing

The game is played in two periods (interpreted as terms in office). Period 1 is divided into four stages, see figure 1. At date $t = 0$, nature draws the incumbent's type θ, the state of the world ω and the signal σ. All types of politicians and voters observe the public signal σ but only type θ_H politicians learn the state of the world ω. At $t = 1$, the incumbent decides which action to implement. At date $t = 2$, the outcome of the policy is realized and learned by all players. At date $t = 3$, the election stage of the game, voters choose between the incumbent and a challenger. The electorate's perception of the challenger (i.e., the probability that the challenger is of high ability) is λ_C, randomly drawn from a distribution on $[0, 1]$.[16] In the second period, following a draw of ω and σ, the appointed politician takes an action. After that, the second-period outcome is realized and publicly observed. Then, the game ends.

Politicians and voters

Voters' task is to decide whether to reelect or replace the incumbent. Their strategy

[16] For simplicity we assume a representative voter in the sense of a pivotal median voter. This assumption implies that politicians act as if confronted with homogeneous voters' beliefs, that is all voters hold the same beliefs about the government. Alternatively, the setup might be interpreted as representing an electorate with three groups of voters: two equally strong partisan groups which always vote for their party's candidate, and a third group of "independent voters" who vote for the candidate which they perceive as more competent, regardless of his affiliation.

consists of a probability distribution over the actions "reelect the incumbent" and "elect the challenger" for each possible combination of signal, action, outcome, and challenger's perceived ability. Voters' payoff equals expected discounted social welfare; their discount factor is given by $\beta \in [0, 1]$. The politician's preferences are given by[17]

$$u = \phi W + (1 - \phi) \Pr[\text{reelection}],$$

where W is social welfare in the current period and $\phi \in (0, 1)$ a weighting factor that determines a politician's relative concern for welfare and reelection.[18] Since there are no reelection concerns in the second term, politician's and voters' objectives are perfectly aligned. All politicians *try to maximize welfare* in the second term but are not equally good at it. Voters' optimal strategy is therefore to elect the candidate they perceive as more competent. Let $\mu(\sigma, a, y)$ denote voters' posterior belief that the politician is of type θ_H given recalled signal σ, policy choice a and realized outcome y.[19] Belief μ is also referred to as the incumbent's *reputation*. The optimal strategy for voters is to reelect the incumbent if and only if $\lambda_C \leq \mu(\sigma, a, y)$. We assume that λ_C is uniformly distributed on $[0, 1]$. Thus, the incumbent's reelection probability is equal to the voters' posterior belief that he is of high ability.

The politician's payoff is his expected utility U. Let α denote a mixed action such that the politician plays a_1 with probability α and a_0 with probability $1 - \alpha$. Hence, expected utility given voters' behavior and the information available to the politician is

$$U(\alpha, \mu, \Psi_\theta) = \alpha\big[\phi E(W(a_1)|\Psi_\theta) + (1 - \phi)E\big(\mu(\sigma, a_1, y)|\Psi_\theta\big)\big] + (1 - \alpha)(1 - \phi)\mu(\sigma, a_0, 0),$$

where Ψ_θ is the politician's (type dependent) information set:

$$\Psi_\theta = \begin{cases} (\omega, \sigma) & \text{for} \quad \theta = \theta_H \\ \sigma & \text{for} \quad \theta = \theta_L. \end{cases}$$

A politician's strategy prescribes a probability $s(\theta, \Psi_\theta)$ of playing a_1 for each type θ and

[17] The assumption that all types of politician care about social welfare, as well as holding office, can be justified, for example, by the fact that politicians are drawn from the population of voters. Thus, they can be expected to consume the same goods as the rest of the electorate. While this formulation follows authors such as Rogoff (1990), Harrington (1993), Canes-Wrone *et al.* (2001) and Majumdar and Mukand (2004), one should note that it is in contrast with other political-agency models which assume that at least some types of politician have their own agenda. This is true for the early work of Barro (1973) and Ferejohn (1986) who use a pure moral hazard framework, but also for more recent contributions who, for the most part, assume that there are both "good" and "bad" types of politician, where good types generally have objectives which are "congruent" with society's whereas bad types favor a special-interest group (see, e.g., Coate and Morris (1995), Maskin and Tirole (2004) and Smart and Sturm (2006)). Reasonable people can disagree whether the congruence or competence dimension is more important in politics. Note, however, that for many political decisions, there are no influential special-interest groups to which politicians might give into.

[18] All players are risk neutral. Notice also that $1 - \phi$ implicitly includes the politician's discount factor.

[19] For rational voters, unlike for hindsight biased voters, the recollected signal σ always coincides with the original signal.

each possible realization of Ψ_θ.[20] We can now define our equilibrium concept, which is a refined version of perfect Bayesian equilibrium (PBE). We omit the voters' strategy from the definition because of its simplicity.

Definition 1 *An equilibrium is a PBE of the game such that* (i) *strategies are optimal given beliefs, i.e.,*

$$\forall\, \theta, \forall\, \Psi_\theta, \quad s^*(\theta, \Psi_\theta) \in \arg\max_{\alpha}\, U(\alpha, \mu, \Psi_\theta),$$

(ii) *beliefs are derived from equilibrium strategies and observed actions using Bayes' rule whenever possible, i.e.,*

$$\mu(\sigma, a, y) = \frac{\lambda_I\ \Pr[\sigma, a, y|\theta_H]}{\lambda_I\ \Pr[\sigma, a, y|\theta_H] + (1-\lambda_I)\Pr[\sigma, a, y|\theta_L]},$$

(iii) *voters hold pessimistic out-of-equilibrium beliefs, i.e.,* $\mu(\sigma, a, y) = 0$ *for any triplet* (σ, a, y) *that is off the equilibrium path, and* (iv) *the D1 criterion is satisfied.*

The two refinements (points (iii) and (iv)) allow us to select a unique equilibrium.[21] In what follows we drop the politician's type from the specification of strategies as this cannot lead to confusion. Thus, we write $s(\omega, \sigma)$ for $s(\theta_H, (\omega, \sigma))$ and $s(\sigma)$ for $s(\theta_L, \sigma)$. We look for an equilibrium in which the high-ability politician always implements the welfare-maximizing policy, that is, he chooses the policy corresponding to the state of nature. Formally, $s(0, \sigma) = 0$ and $s(1, \sigma) = 1$, regardless of σ.[22]

In the next section, we establish the existence of equilibrium under the assumption of perfect rationality on the part of all players and thereafter extend the analysis to the case of hindsight biased voters.

3 Equilibrium with rational voters

With rational voters, the game has two independent subgames, one for each realization of the signal σ, which we will analyze in turn.

[20] In game-theoretic terms, there are, strictly speaking *three* types of politician in this model: two types of high-ability politician – one for each realization of the first-period state of the world – and one type of low-ability politician. We have chosen to label types in more intuitive terms for greater clarity of exposition.

[21] The role of pessimistic beliefs will become clear below. The D1 criterion rules out equilibrium candidates where both types of politician pool on one action (either a_0 or a_1) independent of their information, as we demonstrate in appendix B.

[22] Note, however, that we do not *assume* that high-ability politicians play in such a way; rather, their mechanical play arises as an equilibrium strategy.

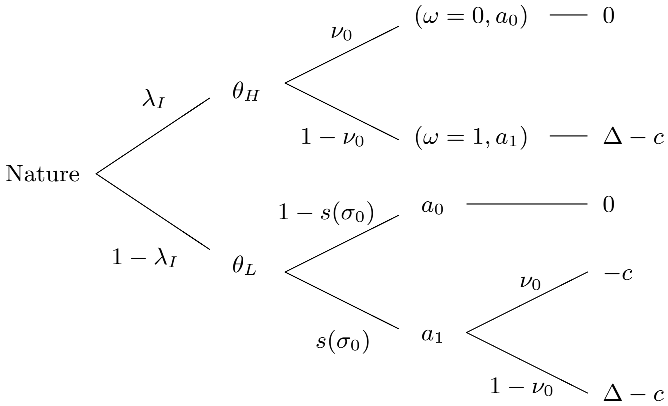

Figure 2: The σ_0 subgame

The σ_0 subgame

Assume $\sigma = \sigma_0$, that is, the signal suggests that the state of the world is $\omega = 0$, and thus that policy a_0 is optimal for welfare. We start by specifying voters' posterior beliefs about the politician's type given an observed triplet (σ_0, a, y) and the low type's strategy $s(\sigma_0)$; see figure 2 which is based on the high type playing mechanically (as explained at the end of section 2). Applying Bayes' rule and pessimistic beliefs, we have, for any $0 \leq s(\sigma_0) \leq 1$,[23]

$$\begin{aligned}
\mu(\sigma_0, a_0, 0) &= \frac{\lambda_I \nu_0}{\lambda_I \nu_0 + (1-\lambda_I)\big(1 - s(\sigma_0)\big)} \\
\mu(\sigma_0, a_1, 0) &= 0 \\
\mu(\sigma_0, a_1, \Delta) &= \frac{\lambda_I}{\lambda_I + (1-\lambda_I) s(\sigma_0)}.
\end{aligned}$$

Denote a low-ability politician's expected utility from playing a_0 by U_0^0, where the superscript stands for the signal σ and the subscript for the chosen policy a. We have

$$\begin{aligned}
U_0^0 &= \phi \cdot 0 + (1-\phi) \cdot \mu(\sigma_0, a_0, 0) \\
&= (1-\phi) \frac{\lambda_I \nu_0}{\lambda_I \nu_0 + (1-\lambda_I)\big(1 - s(\sigma_0)\big)}. \qquad (1)
\end{aligned}$$

[23] If $s(\sigma) = 0$, observing $(\sigma, a_1, 0)$ is an out-of-equilibrium event for which $\mu(\sigma, a_1, 0) = 0$ by pessimistic beliefs. Hence, voters think that any politician whose policy fails must be of low ability. This seems quite natural, not least since it is exactly the belief Bayes' rule specifies for any $s(\sigma) > 0$. Such beliefs would also obtain if there were an additional type of low-ability politician who has a preference for risky actions and thus always plays a_1 regardless of his signal.

Playing a_1 yields $U_1^{0,R}$ given by

$$\begin{aligned} U_1^{0,R} &= \phi\big[(1-\nu_0)\Delta - c\big] + (1-\phi)\Big[\nu_0\,\mu(\sigma_0, a_1, 0) + (1-\nu_0)\,\mu(\sigma_0, a_1, \Delta)\Big] \\ &= \phi\big[(1-\nu_0)\Delta - c\big] + (1-\phi)\frac{(1-\nu_0)\lambda_I}{\lambda_I + (1-\lambda_I)s(\sigma_0)}. \end{aligned} \tag{2}$$

We will categorize equilibria according to the low type's equilibrium strategy: when $s(\sigma_0) = 0$, we will talk about pure-strategy equilibrium, while for $0 < s(\sigma_0) < 1$ (the low type randomizes between a_0 and a_1), we will talk about mixed-strategy equilibrium.[24] For a pure-strategy equilibrium, it must be the case that the low-ability politician prefers a_0 over a_1 even if voters believe that $s(\sigma_0) = 0$, i.e., voters think that any politician who chooses a_1 and succeeds must be of high type. In that case, the low type could fool voters into thinking he is of high ability by successfully implementing the risky policy. If, despite such beliefs, we have $U_0^0 \geq U_1^{0,R}$, then we are in a pure-strategy equilibrium.[25] Otherwise, we are in a mixed-strategy equilibrium where the high type always chooses the "right" policy while the low type randomizes between actions a_0 and a_1. For him to be willing to do so, he must be indifferent between the two policies, that is, both must procure him equal utility in expectation. This requires that voters hold the appropriate beliefs. Lemma 1 describes the conditions under which these equilibria obtain.

Lemma 1 *Suppose voters are rational. There exists a unique equilibrium in the σ_0 subgame characterized by a threshold ν_0^R such that the high-ability politician chooses a_0 when $\omega = 0$ and a_1 when $\omega = 1$, while the low-ability politician always chooses a_0 when $\nu_0 \geq \nu_0^R$ and randomizes between a_0 and a_1 when $\nu_0 < \nu_0^R$. In the latter case, type θ_L's equilibrium probability of playing a_1, $s_R^*(\sigma_0)$, is determined by*

$$\frac{\phi}{1-\phi}\big[(1-\nu_0)\Delta - c\big] = \frac{\lambda_I\nu_0}{\lambda_I\nu_0 + (1-\lambda_I)\big(1-s_R^*(\sigma_0)\big)} - \frac{(1-\nu_0)\,\lambda_I}{\lambda_I + (1-\lambda_I)s_R^*(\sigma_0)}, \tag{3}$$

and $s_R^*(\sigma_0)$ decreases with ν_0. The threshold ν_0^R is defined by

$$\frac{\phi}{1-\phi}\big[(1-\nu_0^R)\Delta - c\big] = \frac{\lambda_I\nu_0^R}{1-\lambda_I(1-\nu_0^R)} - (1-\nu_0^R). \tag{4}$$

[24] In the σ_0 subgame, an equilibrium where the high type always chooses the policy corresponding to the underlying state of nature and where $s(\sigma_0) = 1$ (i.e., the low type always does the opposite of what the signal suggests) can be ruled out. To see why, note first that in this kind of equilibrium, voters believe that any politician who *does* follow the signal and plays a_0 must be of high ability with probability 1. Thus, the low type can increase his reputation by choosing the policy the signal suggests. Moreover, in terms of expected social welfare, the low type is always better off following the signal. Therefore, choosing the "wrong" policy a_1 all the time cannot be an equilibrium. This is shown formally in the proof of Lemma 1.

[25] Technically, it would be inappropriate to talk about a separating equilibrium since from a game-theoretic point of view, the model has three types; see footnote 20. The two high-ability types play pure strategies corresponding to the state of nature. Hence, whatever the action the low type plays, he always "pools" with one of the two high types.

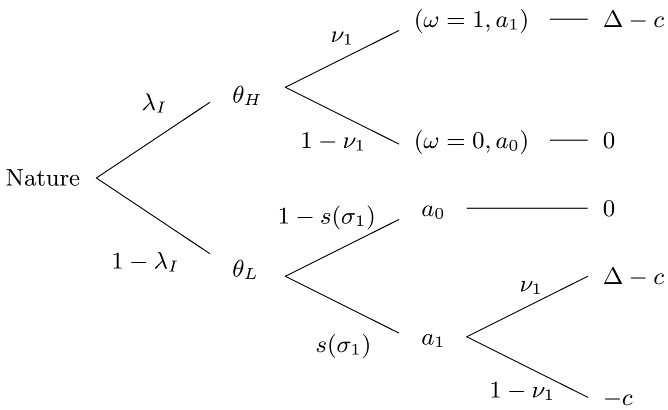

Figure 3: The σ_1 subgame

For this threshold to be located in the admissible interval, i.e., $\nu_0^R \in (1 - c/\Delta, 1)$, it must be the case that $\lambda_I < \frac{c\Delta}{(\Delta-c)^2+c\Delta}$.

Thus, there is a threshold value of signal precision above which we are in a pure-strategy equilibrium where the low type always follows the public signal, and below which we are in a mixed-strategy equilibrium where the low type randomizes. In the mixed-strategy case, the equilibrium probability of playing a_1 decreases with the signal precision.

The condition $\lambda_I < \frac{c\Delta}{(\Delta-c)^2+c\Delta}$ says that the prior probability of the politician being of high ability cannot be too large. Otherwise, the politician's reputation from implementing a_0 is so high that gambling on a_1 is never worthwhile, even when the signal is of low precision. Hence, if this assumption is violated, the politician always follows his signal, regardless of its quality; there is never any randomization by the low type.

The σ_1 subgame

Now assume that $\sigma = \sigma_1$, suggesting that a_1 should be implemented. Posterior beliefs for the σ_1 subgame, depicted in figure 3, are given by

$$\begin{aligned} \mu(\sigma_1, a_0, 0) &= \frac{\lambda_I(1-\nu_1)}{\lambda_I(1-\nu_1) + (1-\lambda_I)\big(1 - s(\sigma_1)\big)} \\ \mu(\sigma_1, a_1, 0) &= 0 \\ \mu(\sigma_1, a_1, \Delta) &= \frac{\lambda_I}{\lambda_I + (1-\lambda_I)s(\sigma_1)}. \end{aligned}$$

When the low-ability politician plays a_1, his expected utility is

$$\begin{aligned} U_1^1 &= \phi(\nu_1\Delta - c) + (1-\phi)\Big[(1-\nu_1)\mu(\sigma_1, a_1, 0) + \nu_1\,\mu(\sigma_1, a_1, \Delta)\Big] \\ &= \phi(\nu_1\Delta - c) + (1-\phi)\frac{\nu_1\lambda_I}{\lambda_I + (1-\lambda_I)s(\sigma_1)}. \end{aligned} \tag{5}$$

Playing a_0 would procure him utility

$$\begin{aligned} U_0^1 &= (1-\phi)\mu(\sigma_1, a_0, 0) \\ &= (1-\phi)\frac{\lambda_I(1-\nu_1)}{\lambda_I(1-\nu_1)+(1-\lambda_I)\big(1-s(\sigma_1)\big)}. \end{aligned} \tag{6}$$

Lemma 2 describes the equilibrium in the σ_1 subgame. It is similar to the one in the σ_0 subgame, with one important qualification: due to the fact that playing a_1 is risky (unlike a_0) and, in case of failure, leads to a bad reputation, there can now be a pure-strategy equilibrium with $s(\sigma_1) = 0$ under certain conditions on parameters; that is, the low-ability politician never follows the signal.[26]

Lemma 2 *Suppose voters are rational. There exists a unique equilibrium in the σ_1 subgame characterized by thresholds $\underline{\nu}_1^R$ and $\overline{\nu}_1^R$ such that the high-ability politician chooses a_0 when $\omega = 0$ and a_1 when $\omega = 1$, while the low-ability politician always chooses a_0 when $\nu_1 \leq \underline{\nu}_1^R$, randomizes between a_0 and a_1 when $\underline{\nu}_1^R < \nu_1 < \overline{\nu}_1^R$, and always chooses a_1 when $\nu_1 \geq \overline{\nu}_1^R$. In the mixed-strategy case, type θ_L's equilibrium probability of playing a_1, $s_R^*(\sigma_1)$, is determined by*

$$\frac{\phi}{1-\phi}(\nu_1\Delta - c) = \frac{\lambda_I(1-\nu_1)}{\lambda_I(1-\nu_1)+(1-\lambda_I)\big(1-s_R^*(\sigma_1)\big)} - \frac{\nu_1\,\lambda_I}{\lambda_I+(1-\lambda_I)s_R^*(\sigma_1)}, \tag{7}$$

and $s_R^*(\sigma_1)$ increases with ν_1. Threshold $\underline{\nu}_1^R$ is defined by

$$\frac{\phi}{1-\phi}(\underline{\nu}_1^R\Delta - c) = \frac{\lambda_I(1-\underline{\nu}_1^R)}{1-\lambda_I\underline{\nu}_1^R} - \underline{\nu}_1^R,$$

having $\underline{\nu}_1^R \in (c/\Delta, 1)$ requires $\lambda_I > \frac{c\Delta}{(\Delta-c)^2+c\Delta}$. Threshold $\overline{\nu}_1^R > \underline{\nu}_1^R$ is defined by

$$\frac{\phi}{1-\phi}(\overline{\nu}_1^R\Delta - c) = 1 - \overline{\nu}_1^R\lambda_I,$$

having $\overline{\nu}_1^R \in (c/\Delta, 1)$ requires $\lambda_I > 1 - \frac{\phi(\Delta-c)}{1-\phi}$.

As in the σ_0 subgame, the low type's equilibrium strategy in the σ_1 subgame depends on signal precision. If $\lambda_I > \frac{c\Delta}{(\Delta-c)^2+c\Delta}$, the low-ability politician always does the opposite of what the signal suggests for very low values of ν_1. For intermediate values of ν_1, the equilibrium has the low-ability politician randomizing. If the condition $\lambda_I > 1 - \frac{\phi(\Delta-c)}{1-\phi}$ is satisfied, the low type always follows the signal for high values of precision. Otherwise, the signal is never sufficiently precise to induce a pure strategy, even when ν_1 is arbitrarily close to 1. Reelection concerns are so strong that the low-ability politician randomizes whatever the signal quality. Figure 4 depicts the low type's equilibrium strategy as a function of the signal precision for both subgames for the case where $1 - \frac{\phi(\Delta-c)}{1-\phi} < \lambda_I < \frac{c\Delta}{(\Delta-c)^2+c\Delta}$.

[26] This cannot happen in the σ_0 subgame, as explained in footnote 24.

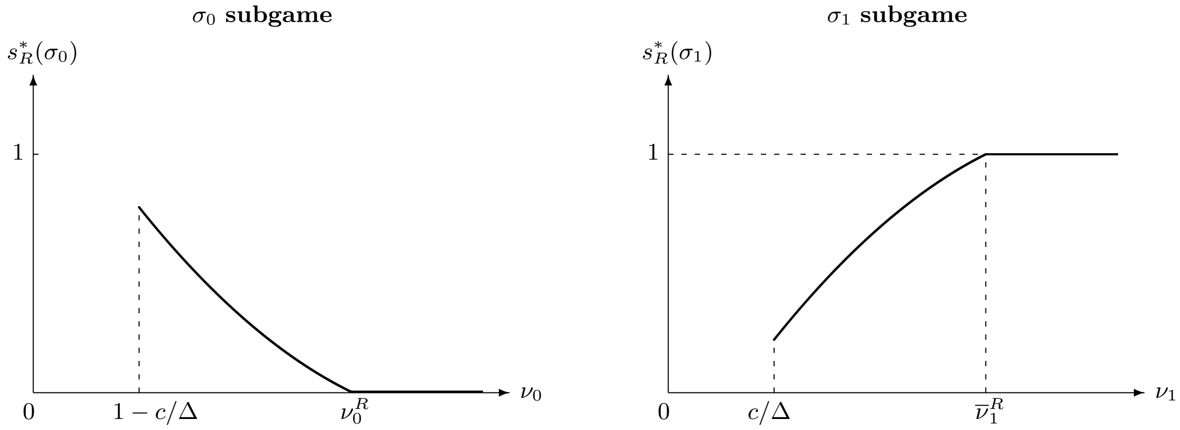

Figure 4: Equilibrium strategies for type θ_L with rational voters

Corollary 1 *The probability that type θ_L plays a_0 (a_1) is greater (smaller) after receiving signal σ_0 than after receiving signal σ_1: $s_R^*(\sigma_0) \leq s_R^*(\sigma_1)$. The inequality is strict if $\lambda_I < \frac{c\Delta}{(\Delta - c)^2 + c\Delta}$.*

This follows directly from Assumption 1 (which implies that $\nu_1 > 1 - \nu_0$) and the monotonicity properties of the equilibrium strategies, noting that $s_R^*(\sigma_0)$ would equal $s_R^*(\sigma_1)$ if $\nu_1 = 1 - \nu_0$. The only case where the inequality is not strict arises when $\nu_1 < \underline{\nu}_1^R$ (for which $\lambda_I > \frac{c\Delta}{(\Delta - c)^2 + c\Delta}$ is necessary) so that $s_R^*(\sigma_0) = s_R^*(\sigma_1) = 0$.

4 Hindsight bias as a discipline device

Voters are hindsight biased if their *recalled prior* about the state of the world diverges from their original prior once they have learned new information about which state truly prevails.[27] More precisely, the bias alters voters' recollection in direction of the publicly observed outcome. It can be formalized using conditional expectations (Camerer *et al.*, 1989):

$$E[E(\omega|\sigma)|\sigma, a, y] = bE(\omega|\sigma, a, y) + (1 - b)E(\omega|\sigma),$$

where $b \in [0, 1]$ measures the degree of hindsight bias. The bias translates into a violation of the law of iterated expectations and as a result the recalled prior belief about ω is located somewhere between the true prior and the posterior probability. However, in the binary model, there are only two possible priors, one for each realization of the signal: $E(\omega|\sigma_0) = 1 - \nu_0$ and $E(\omega|\sigma_1) = \nu_1$. Without restrictions on b, an evaluator's set of recalled prior probabilities

[27] Or, as Rabin (1998, p. 30) puts it, "people exaggerate the degree to which their beliefs before an informative event would be similar to their current beliefs".

would be different from the set of possible original prior probabilities. We therefore choose a formulation of hindsight bias that is consistent with the set of prior beliefs an evaluator may originally hold about the state of the world.

Definition 2 (Hindsight bias with a binary signal) *Hindsight biased voters overestimate the accuracy of their prior belief about the state of the world: If* $\sigma = \sigma_0$ *and they learn that* $\omega = 1$*, they think that ex ante they attached probability* ν_1 *to the state of the world being 1 even though their original signal suggested probability* $1 - \nu_0 < \nu_1$*. Thus, they erroneously believe that the signal was* σ_1 *rather than* σ_0*. Similarly, if* $\sigma = \sigma_1$ *and they learn that* $\omega = 0$*, hindsight biased voters think that their prior was* ν_0 *even when according to their signal it was* $1 - \nu_1 < \nu_0$*; they think that their signal was* σ_0 *rather than* σ_1*.*

The safe policy a_0 is uninformative for voters in terms of its outcome.[28] The recollection of prior probabilities is only altered when new information about the state of the world is revealed, i.e., when the risky policy a_1 is implemented. After policy a_1, voters learn the state of the world with certainty: posterior beliefs over ω are $E(\omega|\sigma, a_1, \Delta) = 1$ and $E(\omega|\sigma, a_1, 0) = 0$.[29] Table 1 summarizes biased voters' recalled signal as a function of outcome y.

	Original signal	
(after a_1)	$\sigma = \sigma_0$	$\sigma = \sigma_1$
$y = 0$	σ_0	σ_0
$y = \Delta$	σ_1	σ_1

Table 1: Biased recollection of original signal σ

In solving for the equilibrium of the game with hindsight biased voters we assume that politicians anticipate voters' hindsight bias and maintain the assumption that voters know the politician's equilibrium strategy.[30] One generally has to consider both subgames. It turns out, however, that the σ_1 subgame is unaffected by hindsight bias. In case of failure of a risky policy, the bias changes an evaluator's posterior belief from $\mu(\sigma_1, a_1, 0)$ to $\mu^B(\sigma_1, a_1, 0) = \mu(\sigma_0, a_1, 0)$. But these posteriors are both equal to zero, whatever the low type's equilibrium strategy, because of the assumption of pessimistic beliefs. Therefore, when $\sigma = \sigma_1$, the

[28] Action a_0 always leads to outcome $y = 0$ and in this case signals are not distorted by hindsight biased evaluators because $E(\omega|\sigma_0, a_0, 0) = E(\omega|\sigma_0)$ and $E(\omega|\sigma_1, a_0, 0) = E(\omega|\sigma_1)$.

[29] We can then identify the parameter b which implicitly underlies our setup as $\nu_1 = b \cdot 1 + (1-b)(1-\nu_0) \iff b = (\nu_1 + \nu_0 - 1)/\nu_0$ in the σ_0 subgame, and $1 - \nu_0 = b \cdot 0 + (1-b)\nu_1 \iff b = (\nu_1 + \nu_0 - 1)/\nu_1$ in the σ_1 subgame.

[30] See section 4.2 for a discussion concerning players' bias awareness and its implications.

hindsight biased equilibrium strategies, $s_B^*(\sigma_1)$ for the low type and $s_B^*(\omega, \sigma_1)$ for the high type, are the same as in the equilibrium with rational voters, see Lemma 2.[31] This only leaves the subgame following signal σ_0 to be analyzed.

Unlike under rational evaluation, the σ_0 subgame under biased evaluation is not independent of the σ_1 subgame. This is because the evaluator calculates his posterior about the politician's type with a biased prior probability. As will become clear, posterior beliefs – and hence, equilibrium strategies – in the σ_0 subgame depend on the equilibrium strategy $s_R^*(\sigma_1)$ of the σ_1 subgame.[32] For the remainder of our analysis, we make an assumption that ensures that hindsight bias is relevant.

Assumption 2 *The following condition on parameters holds:*

$$\lambda_I < \frac{c\Delta}{(\Delta - c)^2 + c\Delta}.$$

If this assumption is not satisfied, the low type's equilibrium strategy is $s_R^*(\sigma_0) = 0$ regardless of the quality of σ_0, and hindsight bias has no impact whatsoever.

4.1 Equilibrium with biased voters

Assume $\sigma = \sigma_0$, so the signal implies (by Assumption 1) that action a_0 is optimal in terms of welfare. The voters' posterior beliefs about the politician's type depend on the observed event (a, y), the recalled signal and the politician's strategy. As before, we start from the premise that the high type chooses the optimal policy given the state of the world. We then show that this is indeed an equilibrium. By Bayes' rule and the assumption of pessimistic beliefs, biased voters' posterior beliefs are, for any $0 \leq s(\sigma_0) \leq 1$,

$$\begin{aligned} \mu(\sigma_0, a_0, 0) &= \frac{\lambda_I \nu_0}{\lambda_I \nu_0 + (1 - \lambda_I)\big(1 - s(\sigma_0)\big)} \\ \mu(\sigma_0, a_1, 0) &= 0 \\ \mu^B(\sigma_0, a_1, \Delta) \;=\; \mu(\sigma_1, a_1, \Delta) &= \frac{\lambda_I}{\lambda_I + (1 - \lambda_I) s_R^*(\sigma_1)}. \end{aligned}$$

The effects of hindsight bias come directly into play at the posterior belief for realization (σ_0, a_1, Δ). Upon observing outcome Δ, biased voters learn that the state of the world is $\omega = 1$. They then distort their recollection of the prior belief, which was based on the original signal σ_0, toward their ex post information by wrongly believing that the signal had been σ_1, therefore $\mu^B(\sigma_0, a_1, \Delta) = \mu(\sigma_1, a_1, \Delta)$. The expected utility for the θ_L type if he plays action a_0 is

$$U_0^0 = \phi \cdot 0 + (1 - \phi) \cdot \mu(\sigma_0, a_0, 0), \tag{8}$$

[31] In other words, equilibrium strategies in the σ_1 subgame are $s_B^*(\sigma_1) = s_R^*(\sigma_1)$ for a low type and $s_B^*(\omega, \sigma_1) = s_R^*(\omega, \sigma_1)$ for a high type.

[32] The reverse is not true since the σ_1 subgame is unaffected by hindsight bias, as mentioned above.

while the deviating action a_1 yields

$$U_1^{0,B} = \phi\big[(1-\nu_0)\Delta - c\big] + (1-\phi)\big[\nu_0\,\mu(\sigma_0, a_1, 0) + (1-\nu_0)\,\mu^B(\sigma_0, a_1, \Delta)\big]. \tag{9}$$

For a pure-strategy equilibrium, a θ_L type must prefer a_0 over a_1, that is, $U_0^0 \geq U_1^{0,B}$, even if voters believe $s(\sigma_0) = 0$. Otherwise we are in a mixed-strategy equilibrium in which the low type randomizes between a_0 and a_1 such that $U_0^0 = U_1^{0,B}$, with voters correctly anticipating equilibrium strategies. Lemma 3 describes equilibrium behavior with hindsight biased voters in the σ_0 subgame.

Lemma 3 *Suppose voters are hindsight biased. There exists a unique equilibrium in the σ_0 subgame characterized by a threshold ν_0^B such that the high-ability politician chooses a_0 when $\omega = 0$ and a_1 when $\omega = 1$, while the low-ability politician always chooses a_0 when $\nu_0 \geq \nu_0^B$ and randomizes between a_0 and a_1 when $\nu_0 < \nu_0^B$. In the latter case, type θ_L's equilibrium probability of playing a_1, $s_B^*(\sigma_0)$, is determined by*

$$\frac{\phi}{1-\phi}\big[(1-\nu_0)\Delta - c\big] = \frac{\lambda_I \nu_0}{\lambda_I \nu_0 + (1-\lambda_I)\big(1 - s_B^*(\sigma_0)\big)} - \frac{(1-\nu_0)\lambda_I}{\lambda_I + (1-\lambda_I)s_R^*(\sigma_1)}, \tag{10}$$

and $s_B^*(\sigma_0)$ decreases with ν_0. The threshold ν_0^B is defined by

$$\frac{\phi}{1-\phi}\big[(1-\nu_0^B)\Delta - c\big] = \frac{\lambda_I \nu_0^B}{1 - \lambda_I(1-\nu_0^B)} - \frac{(1-\nu_0^B)\lambda_I}{\lambda_I + (1-\lambda_I)s_R^*(\sigma_1)}. \tag{11}$$

For ν_0^B to be in the admissible range, i.e., $\nu_0^B \in (1 - c/\Delta, 1)$, it must be the case that $\lambda_I < \frac{(\Delta-c)\Delta\big(1-s_R^*(\sigma_1)\big)+2c\Delta-\Delta^2}{(\Delta-c)\Delta\big(1-s_R^*(\sigma_1)\big)+c^2}$.

As in the case of rational voters, the low type's behavior depends on the precision of σ_0: for high values of ν_0, the low-ability politician implements a_0, for low values he randomizes between a_0 and a_1. However, the cutoff and the equilibrium probability of playing a_1 for a given ν_0 are not the same as with rational voters. The following proposition assesses the impact of hindsight bias on the low-ability politician's decision making.

Proposition 1 *For any $\nu_0 < \nu_0^R$, hindsight bias strictly improves the low-ability politician's discipline: $s_B^*(\sigma_0) < s_R^*(\sigma_0)$.*

Proposition 1 means that, in terms of first-period social welfare, voters benefit from being hindsight biased. The intuition for this result is the following. Biased voters are less easily impressed by deviating behavior. They consider the outcome Δ in retrospect more predictable than it actually was, and therefore think that playing a_1 was the "obvious" policy choice even for a low-ability politician. In other words, they exaggerate the likelihood that the observed

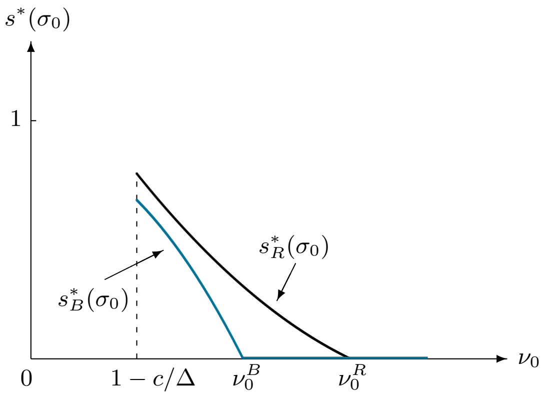

Figure 5: Rational and biased equilibrium strategies for type θ_L

event came from a low type, which reduces their esteem for the incumbent.[33] Anticipating voters' bias, a low-ability politician knows that he has relatively little to gain from deviating to policy a_1, and accordingly, will do so less often. This is illustrated in figure 5, in which equilibrium strategies for a low-ability decision maker in the σ_0 subgame under rational and biased evaluation are contrasted.

4.2 Discussion

We discuss the psychological foundation of hindsight bias and then look at voters' bias awareness and robustness to alternative assumptions.

Imperfect memory or self-serving bias

In psychology, judgment biases are explained by either motivational or cognitive theories. Motivational theories rationalize the existence of a judgment bias by a deliberate but often subconscious choice of the decision maker who may derive a (psychological) benefit from it.[34] In cognitive theories a judgment bias is attributed to information processing effects. Even though motivational aspects may reinforce hindsight bias, research in psychology seems to identify cognitive effects as its main source (Hawkins and Hastie, 1990; Hoffrage *et al.*, 2000). Therefore, we assume that self-serving effects play no role for voters in our problem and rather model hindsight bias as a by-product of knowledge updating after outcome information has

[33] Voters wrongly recall their prior as σ_1 and therefore think the low type's equilibrium strategy is $s_R^*(\sigma_1)$. By Corollary 1, a low-ability politician is more likely to play a_1 after σ_1 than after σ_0.

[34] Hindsight bias may be self-serving for various reasons. An individual may want to distort his recollection of the original prior because he derives a benefit from appearing smart in front of others or himself. Moreover, suppression of changes in probability assessments over time decreases an individual's perception of uncertainty in the world.

been received.[35] According to cognitive theories, a biased evaluator's default memory consists of current, up to date probability estimates but does not stock all previously formed prior probabilities because new information about the true state of the world destroys or reduces their accessibility.[36] Thus, a subject has to follow a mental strategy to *reconstruct* the original prior from the default information he holds ex post; this is what Hawkins and Hastie (1990, p. 321) call "reconstruction of the prior judgment by 'rejudging' the outcome".

Awareness

To solve the model with hindsight bias we use a variation of perfect Bayesian equilibrium (PBE) which captures memory imperfections on the part of evaluators. Our approach relies on the assumption that politicians anticipate voters' hindsight bias and that voters correctly predict the equilibrium strategy used by a politician for each realization of the signal.[37] However, voters' posterior beliefs are based on the strategy associated with the signal they recall, rather than the original signal. Therefore they may hold incorrect posterior beliefs, but the mistake stems solely from the erroneous recollection of the prior and not from wrong expectations about strategies.

Notice also that voters are *naive*, that is ex post unaware of their bias: they are certain that they correctly remember the signal. Voters' imperfect recollection process thus hinders conscious learning and implies a reduction of surprises of any kind. A natural extension is to depart from the assumption of naïveté. Suppose evaluators are aware of their possibly distorted recollection of the ex ante signal at the voting stage, i.e., they are *sophisticated*.[38] Denote the probability that the signal corresponds to the underlying state of nature by $x \equiv Pr[\sigma_\omega|\omega]$ with $\omega \in \{0,1\}$, and the probability that the state of the world is 0 by $\pi \in (0,1)$.[39] The *reliability* of voters' memory is the probability that the recollected signal corresponds to the original one. We denote the reliability of a recalled signal σ_1 (given event (a_1, Δ)) by r.

[35] An alternative modeling approach, in the spirit of Bénabou and Tirole (2006), would give the evaluator an opportunity for memory manipulation, or (motivated) manipulation of own beliefs about the world.

[36] Hoffrage *et al.* (2000) emphasize the fact that for subjects with capacity-constrained memory, holding current information in memory is, for general tasks, more important and accurate than remembering past prior probabilities which are, by definition, based on outdated information. For a memory-based model of bounded rationality in economic theory see Mullainathan (2002); there, as in our model, a decision maker takes the recalled history of signals as the true history.

[37] In fact, it does not matter whether voters calculate strategies themselves or whether they learn them from experience or other sources, although the interpretations are different. In the former case voters are aware of their hindsight bias ex ante, but not ex post. Interestingly, such behavior is consistent with a recent study by DellaVigna and Kaplan (2007) which shows that even voters who are aware of biased media coverage on television do not fully subtract the broadcaster's political bias in their voting decision. In the latter case they are ignorant of their bias, or naive, at all stages of the game.

[38] For a model where individuals forget or repress information, but are aware of the deficiencies of their memory, see Bénabou and Tirole (2002).

[39] Since $\nu_0 = \frac{\pi x}{\pi x+(1-\pi)(1-x)}$ and $\nu_1 = \frac{(1-\pi)x}{(1-\pi)x+\pi(1-x)}$, x and π must satisfy $\frac{x}{1-x} > \frac{(1-\pi)(\Delta-c)}{\pi c} > \frac{1-x}{x}$ to be consistent with Assumption 1.

We then have

$$r = \frac{x\ (\lambda_I + (1-\lambda_I)s_R^*(\sigma_1))}{x\ (\lambda_I + (1-\lambda_I)s_R^*(\sigma_1)) + (1-x)\ (\lambda + (1-\lambda)s_S^*(\sigma_0))}, \tag{12}$$

where $s_S^*(\sigma_0)$ is the equilibrium strategy under sophisticated evaluation. A θ_L type's payoff from playing a_1 after σ_0 becomes

$$U_1^{0,S} = \phi[(1-\nu_0)\Delta - c] + (1-\phi)(1-\nu_0)\Big[r\,\mu(\sigma_1, a_1, \Delta) + (1-r)\,\mu(\sigma_0, a_1, \Delta)\Big]. \tag{13}$$

In the next proposition, we compare the low-ability politician's equilibrium strategy under sophistication to the benchmark results with rational and naive hindsight biased voters.

Proposition 2 *Sophistication lowers the impact of hindsight bias but does not eliminate its disciplining effect on low-ability politicians:* $s_B^*(\sigma_0) \leq s_S^*(\sigma_0) < s_R^*(\sigma_0)$ *for any* $\nu_0 < \nu_0^R$.

Thus, the positive effect of hindsight bias on discipline demonstrated in Proposition 1 is robust to different awareness assumptions. Intuitively, since a politician's reputation with sophisticated voters (last term in (13)) is a convex combination of those with naive hindsight biased and rational voters, the essence of the previous analysis remains intact, although the magnitude of the effect may be reduced.

Robustness to alternative assumptions

So far we have assumed that the entire electorate is hindsight biased. What happens if a fraction ρ of the population is rational? In case of a successful gamble (a_1, Δ) and an ex ante signal σ_0, a fraction ρ of voters has belief $\mu(\sigma_0, a_1, \Delta)$ while a fraction $1-\rho$ has belief $\mu(\sigma_1, a_1, \Delta)$. Hindsight biased voters' opinion about the election decision differs from that of rational voters only if the challenger's perceived ability λ_C satisfies $\mu(\sigma_1, a_1, \Delta) < \lambda_C < \mu(\sigma_0, a_1, \Delta)$. In that case, the rational fraction of the population votes for the incumbent, who they believe is more competent, while the hindsight biased population disagrees and votes for the challenger. Thus, if $\rho > 1/2$ rationality decides the election, the incumbent is confirmed in office and the outcome of the election is unaffected by hindsight bias. The reverse is true for $\rho < 1/2$.

In the model analyzed above the signal can only take two values. An alternative assumption is to have a continuous instead of a binary signal.[40] Hindsight bias could then be modeled as a shift of the recollected signal towards the outcome that actually occurred. For the event $(a_1, 0)$, the analysis would not be modified because voters' posterior would be zero regardless of the signal. However, in the event (a_1, Δ) a continuous signal would complicate

[40] In terms of the welfare-maximizing policy, there would be a cutoff signal below which policy a_0 is optimal and above which policy a_1 is.

the analysis considerably. Unlike in the binary case where only σ_0 is distorted while σ_1 is correctly remembered (see table 1), hindsight bias would affect voters' recollection for all possible realizations of the ex ante signal (except at the upper bound of the distribution). This implies that hindsight bias would not be unambiguously beneficial in terms of discipline; the overall effect on discipline would depend on the distribution of the signal. Our results would remain valid for some, but not all, distributional assumptions.

5 Selection and welfare

Proposition 1 shows that hindsight bias improves the discipline of the (low-ability) incumbent. Therefore, the effect of hindsight bias on voters' first-period welfare is unambiguously positive. To make a general statement about the impact of hindsight bias on welfare, however, we must also take into account second-period welfare. This means we have to investigate the effect on selection: Because in the second period, the politician always implements the policy that, according to his information, is best, voters are weakly better off if a high-ability politician is in office.

We can focus on the case where the first-period signal is σ_0, since hindsight bias changes nothing when the signal is σ_1. The effect of hindsight bias on selection works through two channels. The first is that voters sometimes have erroneous posteriors, so that they don't always elect the politician that is truly more able (in expected terms); this clearly is bad for welfare. Hindsight bias essentially blurs the two major elements voters initially set out to distinguish in their evaluation, namely skill and luck of the decision maker. The second is more indirect: since the anticipation of voters' hindsight bias changes the low-ability politician's behavior, the inferences that can be drawn from a given event are also modified. The welfare implications of this second effect are more complex. Hindsight bias increases the low type's equilibrium probability of playing a_0 after receiving σ_0. This decreases the posterior probability of a high type after observing a_0, while the posterior probability of a high type after observing (a_1, Δ) increases.

Since, in general, both of these effects are at work, the selection effects of hindsight bias are ambiguous. However, it follows from the above discussion that, if we fix $s(\sigma_0)$, only the first effect operates. In that case, hindsight bias is detrimental to selection. This allows us to derive a polar case for which hindsight biased selection is dominated by rational selection.

Formally, denote the expected second-period welfare when a low (high) type is in office by w_L (w_H).[41] Clearly, $w_H > w_L$. Let $w(z) = zw_H + (1-z)w_L$ be the expected second-

[41] Although of no relevance for the analysis, these expectations can be calculated. Since in the second period, the low-ability politician always follows his signal, we have: $w_H = (1-\pi)(\Delta - c)$ and $w_L = (1-\pi)x(\Delta - c) - \pi(1-x)c$.

period welfare when the probability that the politician has high ability is z. Therefore, expected second-period welfare when an incumbent with reputation μ is reelected is $w(\mu)$, while expected welfare when a challenger of perceived ability λ_C is elected is $w(\lambda_C)$. For notational simplicity, we will adopt the following conventions throughout this section:

$$\begin{aligned}\mu_0(s) &\equiv \frac{\lambda_I \nu_0}{\lambda_I \nu_0 + (1-\lambda_I)(1-s)} \\ \mu_{1\Delta}(s) &\equiv \frac{\lambda_I}{\lambda_I + (1-\lambda_I)s}\end{aligned}$$

where $\mu_0(s)$ is voters' posterior belief that the politician is competent given that he chose a_0 after observing σ_0 and that the low-type's strategy is s, while $\mu_{1\Delta}(s)$ is the posterior belief after event (a_1, Δ).[42]

After the first-period signal σ_0 (but before policy choice and outcome realization), the ex ante expected second-period welfare with rational voters, as a function of the low type's strategy s, can be written as

$$\begin{aligned}EW_R(s) = \big(\nu_0\lambda_I + (1-\lambda_I)(1-s)\big)&\left[\int_0^{\mu_0(s)} w\big(\mu_0(s)\big)\mathrm{d}\lambda_C + \int_{\mu_0(s)}^1 w\big(\lambda_C\big)\mathrm{d}\lambda_C\right] + \\ + \nu_0(1-\lambda_I)s\int_0^1 w\big(\lambda_C\big)\mathrm{d}\lambda_C &+ \\ + (1-\nu_0)\big(\lambda_I + (1-\lambda_I)s\big)&\left[\int_0^{\mu_{1\Delta}(s)} w\big(\mu_{1\Delta}(s)\big)\mathrm{d}\lambda_C + \int_{\mu_{1\Delta}(s)}^1 w\big(\lambda_C\big)\mathrm{d}\lambda_C\right].\end{aligned}$$

In words, ex ante expected second-period welfare is the weighted sum of welfare in each of the three possible events $(a_0, 0)$, $(a_1, 0)$ and (a_1, Δ), taking expectations over the challenger's ability. Turning to the hindsight biased case, ex ante welfare as a function of the low type's strategy is

$$\begin{aligned}EW_B(s) = \big(\nu_0\lambda_I + (1-\lambda_I)(1-s)\big)&\left[\int_0^{\mu_0(s)} w\big(\mu_0(s)\big)\mathrm{d}\lambda_C + \int_{\mu_0(s)}^1 w\big(\lambda_C\big)\mathrm{d}\lambda_C\right] + \\ + \nu_0(1-\lambda_I)s\int_0^1 w\big(\lambda_C\big)\mathrm{d}\lambda_C &+ \\ + (1-\nu_0)\big(\lambda_I + (1-\lambda_I)s\big)&\left[\int_0^{\mu_{1\Delta}(s_R^*(\sigma_1))} w\big(\mu_{1\Delta}(s)\big)\mathrm{d}\lambda_C + \int_{\mu_{1\Delta}(s_R^*(\sigma_1))}^1 w\big(\lambda_C\big)\mathrm{d}\lambda_C\right].\end{aligned}$$

This expression reveals the effect of hindsight bias on selection: a politician who succeeds with a risky policy is reelected with probability $\mu_{1\Delta}(s_R^*(\sigma_1))$, even though the actual posterior probability of him being a high type is $\mu_{1\Delta}(s)$.

[42] We could define $\mu_{10}(s)$ analogously. Recall, though, that the posterior belief for the event $(a_1, 0)$ is zero for any s.

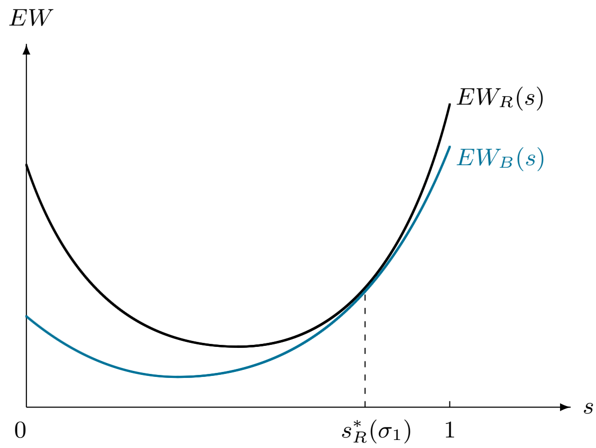

Figure 6: Expected second-period welfare: Rational vs biased voters

Lemma 4 *For a given s, rational voters can expect to be better off in the second period than hindsight biased voters:*

$$EW_R(s) \geq EW_B(s) \quad \forall s \in [0,1]$$

with strict inequality for $s \neq s_R^*(\sigma_1)$.

Lemma 4, illustrated in figure 6, confirms the above-mentioned intuition according to which, for a given s, only the negative effect of hindsight bias is at work, which is to distort posterior beliefs and thereby lead to wrong voting decisions. The following proposition exploits this result to show that hindsight bias is detrimental to selection for particular values of signal precision.

Proposition 3 *In terms of selection, voters are better off being rational rather than hindsight biased for any $\nu_0 \geq \nu_0^R$ and for a nonempty set of values of ν_0 below ν_0^R.*

Thus, hindsight bias is bad for voters' expected second-period welfare at least when ν_0 is in the vicinity of ν_0^R. When moving further below ν_0^R, the effect of hindsight bias on selection is ambiguous and depends on parameters; hindsight bias can sometimes even improve selection. But even for those cases where the bias harms voters in terms of second-period welfare, this effect has to be contrasted with the positive effect on discipline (and thus first-period welfare).

Proposition 4 *Whatever the effect of hindsight bias on selection, there exists a discount factor $\hat{\beta} \in (0,1]$ such that hindsight bias improves overall welfare for any $\beta \leq \hat{\beta}$.*

According to this proposition, hindsight bias can be welfare-enhancing regardless of what happens in the second period: because voters discount the future, discipline is more important than selection for a sufficiently low discount factor.

Finally, we examine the reelection chances for each type of politician (low-ability and high-ability). This will allow us to assess the impact of hindsight bias on political turnover, defined as the rate of replacement of the politician holding office. The following lemma compares the posterior beliefs held by rational and hindsight biased voters in the σ_0 subgame upon observing an action and outcome.

Lemma 5 *If $\nu_0 < \nu_0^R$, the following inequalities hold for the politician's reputation:*

$$\mu_0\big(s_R^*(\sigma_0)\big) > \mu_0\big(s_B^*(\sigma_0)\big) \tag{14}$$

$$\mu_{10}\big(s_R^*(\sigma_0)\big) = \mu_{10}\big(s_B^*(\sigma_0)\big) = 0 \tag{15}$$

$$\mu_{1\Delta}\big(s_R^*(\sigma_0)\big) > \mu_{1\Delta}\big(s_R^*(\sigma_1)\big). \tag{16}$$

Given Lemma 5, the analysis of the politician's reelection chances becomes straightforward. Consider first the θ_L-type politician. With rational voters, his ex ante probability of reelection when the signal is σ_0, which we denote $\mathcal{R}_L^R$, is given by

$$\mathcal{R}_L^R = (1 - s_R^*(\sigma_0))\,\mu_0\big(s_R^*(\sigma_0)\big) + s_R^*(\sigma_0)(1-\nu_0)\,\mu_{1\Delta}\big(s_R^*(\sigma_0)\big),$$

while in the case of hindsight biased voters, it is

$$\mathcal{R}_L^B = (1 - s_B^*(\sigma_0))\,\mu_0\big(s_B^*(\sigma_0)\big) + s_B^*(\sigma_0)\,(1-\nu_0)\,\mu_{1\Delta}\big(s_R^*(\sigma_1)\big).$$

Turning to the θ_H type, his probability of reelection with rational voters is

$$\mathcal{R}_H^R = \nu_0\,\mu_0\big(s_R^*(\sigma_0)\big) + (1-\nu_0)\mu_{1\Delta}\big(s_R^*(\sigma_0)\big), \tag{17}$$

while with hindsight biased voters, it is

$$\mathcal{R}_H^B = \nu_0\,\mu_0\big(s_B^*(\sigma_0)\big) + (1-\nu_0)\mu_{1\Delta}\big(s_R^*(\sigma_1)\big). \tag{18}$$

The next proposition compares the politician's reelection prospects in the presence of rational and hindsight biased voters.

Proposition 5 *Suppose $\nu_0 < \nu_0^R$. Both low-ability and high-ability politicians are less likely to be reelected when facing hindsight biased voters than when facing rational voters: $\mathcal{R}_L^B < \mathcal{R}_L^R$ and $\mathcal{R}_H^B < \mathcal{R}_H^R$. Thus, hindsight bias increases political turnover.*

Our model predicts that political turnover is larger when voters are hindsight biased (because both low and high type are less likely to be reelected). This result follows directly from the fact that, irrespective of the outcome, rational voters always have a (weakly) higher opinion of the politician than hindsight biased ones; see Lemma 5. The result is in line with conventional wisdom which holds that, when evaluating somebody else's performance, an individual suffering from hindsight bias gives less credit than is due in case of success, and more blame than is warranted in case of failure.

6 Conclusion

We have constructed a political-agency model where voters exhibit a cognitive deficiency known as hindsight bias: after the uncertainty about an event is resolved, they consider the realized outcome more foreseeable than it actually was. In the model, voters have to evaluate the incumbent politician in order to decide whether to reelect him or replace him with a challenger. Politicians are assumed to differ in ability, where ability corresponds to the quality of their information about the welfare-maximizing policy. High-ability politicians are better informed than low-ability politicians and voters. In this setup, low-ability politicians have incentives to disregard public information on what the optimal policy is in order to appear to have superior private information. We have shown that hindsight bias on the part of voters can act as a discipline device. This is because hindsight biased voters are less easily impressed by a successful gamble – they think it was the obvious choice to make from the outset, even if the available information had suggested otherwise. Therefore, they give an incumbent who succeeds with a risky policy in spite of public pessimism less credit than rational voters who perfectly recall their prior. Anticipating this, low-ability politicians are less likely to deviate from the welfare-maximizing policy.

The disciplining effect of hindsight bias is unambiguously beneficial for voters' first-period welfare. However, an overall welfare assessment also has to take into account the second (i.e., post-election) period. Hindsight biased voters' evaluation of an incumbent's competence is less precise than that of rational voters. There is some range of signal precision for which this effect dominates any potential offsetting benefits for selection that hindsight bias may generate through its effect on the low type's behavior. In this case, hindsight bias does not serve voters well in terms of second-period welfare. Nevertheless, as far as *overall* welfare is concerned, hindsight bias can be welfare-enhancing irrespective of its selection effects if voters discount future payoffs at a sufficiently high rate. A final result, in line with conventional wisdom about hindsight bias, is that both the low- and the high-ability politician are less likely to be reelected when voters are hindsight biased than when they are rational, i.e., political turnover is increased.

Due to the multidimensional nature of real-life politics, proving the empirical relevance of hindsight bias is a difficult task. We provide here some anecdotal evidence from an episode of U.S. politics for which public opinion is particularly well-documented: the Gulf War of 1991 (Mueller, 1994). A case can be made that hindsight bias contributed to George Bush Sr.'s defeat in the 1992 presidential election. Bush had initiated military action in response to the Iraqi invasion of Kuwait, an endeavor of considerable political risk, and come away with what observers unanimously viewed as a huge success. With Bayesian voters, Bush's success in Iraq should have shown up positively in his foreign-policy approval rate. In the immediate post-war euphoria, his approval rate did indeed go up. However, by April of 1992, approval for the President's foreign policy was back to (or even slightly below) its pre-war level.[43] This seems to indicate that voters, in retrospect, did not give Bush much credit for the successful operation in the Gulf, an interpretation which is corroborated by opinion polls bearing out voters' hindsight bias (see the introductory quote by Mueller (1994)). If the use of military force seemed like an obvious choice to voters, the war will not have had much impact on Bush's perceived competence, and therefore not have helped him much in the November 1992 election, which he lost to Bill Clinton.[44]

Our analysis may be applicable to problems beyond political economy, for example, promotion decisions in organizations (which, much like elections, do not follow rules set forth in an explicit ex ante contract). Consider a human resource department that has to decide whether to promote an employee from inside the firm, whose actions and performance have been observed, or to hire an outsider for the job. In a firm, there typically is some amount of public information concerning the way an employee is supposed to handle his task (in terms of the model, what the right choice of action is), but employees may also have superior information on their specific assignment. Our model would predict that, if anticipated, hindsight bias on the part of the human resource manager (or decision committee) may prevent low-ability employees from deviating to suboptimal actions in order to appear smart, but would not necessarily help in choosing the right candidate.

We close by noting that, with the benefit of hindsight, all of our results are, of course, obvious.

[43] See Mueller (1994, table 3); approval rates for Bush's handling of the situation in the Middle East follow the same pattern. The gradual decline in the approval rate is consistent with experimental evidence according to which hindsight bias increases over time (Bryant and Brockway, 1997).

[44] There is some question as to whether the war in Iraq was an important factor in voters' election decision. While as late as September 1992, almost 70% of likely voters indeed said that the war was important (Mueller, 1994, table 282), political scientists have come to view the election as largely decided by issues other than foreign policy. Notice, however, that this is not inconsistent with hindsight bias having influenced the election. In fact, if voters thought that the decision to use military force was a "no-brainer", its favorable outcome should not have played much of a role in their updating of the President's perceived competence, compared to other, seemingly more informative issues, such as his handling of the economy.

Appendix

A Proofs

Proof of Lemma 1:

The proof proceeds as follows. First, taking as given that the high-ability politician plays mechanically as stated in the lemma, we derive the conditions for pure-strategy equilibria with either $s^*_R(\sigma_0) = 0$ or $s^*_R(\sigma_0) = 1$. We show that an equilibrium with $s^*_R(\sigma_0) = 1$ is impossible under Assumption 1; we then prove the existence of a threshold ν_0^R as described in the lemma and establish when it is within the set of values that are admissible for ν_0 under Assumption 1. Second, we show that the high-ability politician indeed finds it optimal to follow the equilibrium strategy.

A necessary and sufficient condition for a pure-strategy equilibrium where $s(\sigma_0) = 0$ is that type θ_L prefers a_0 even though voters believe that he will always follow the signal. This condition is obtained by evaluating U_0^0 and $U_1^{0,R}$ at $s(\sigma_0) = 0$, that is

$$\phi[(1-\nu_0)\Delta - c] \leq (1-\phi)\left[\frac{\lambda_I \nu_0}{1-\lambda_I(1-\nu_0)} - (1-\nu_0)\right]. \tag{19}$$

A pure-strategy equilibrium with $s(\sigma_0) = 1$ would require that the low type prefer a_1 even though he could fool voters about his type by playing a_0. The associated condition is obtained by evaluating U_0^0 and $U_1^{0,R}$ at $s(\sigma_0) = 1$:

$$\phi[(1-\nu_0)\Delta - c] \geq (1-\phi)\left[1 - \lambda_I(1-\nu_0)\right]. \tag{20}$$

If neither (19) nor (20) hold, we are in a mixed-strategy equilibrium where the low type randomizes between actions a_0 and a_1.

As $\nu_0 \to 1$, condition (19) is always satisfied, while as $\nu_0 \to 0$, it can never be satisfied. Together with the monotonicity of both left- and right-hand side of (19) with respect to ν_0, this proves that there exists a unique threshold ν_0^R as stated in the lemma. However, condition (20) cannot be satisfied for any $\nu_0 > 1 - c/\Delta$; thus, there is no equilibrium with $s(\sigma_0) = 1$. For the threshold ν_0^R to be greater than $1 - c/\Delta$, the limit of the right-hand side of (19) must be negative as $\nu_0 \to 1 - c/\Delta$ (the left-hand side tends to zero). That is,

$$\lim_{\nu_0 \to 1-c/\Delta} \frac{\lambda_I \nu_0}{1-\lambda_I(1-\nu_0)} - (1-\nu_0) < 0$$

$$\Leftrightarrow \lambda_I < \frac{c\Delta}{(\Delta - c)^2 + c\Delta}.$$

For values below ν_0^R, the equilibrium is in mixed strategies. For the type θ_L politician to be willing to randomize, voters' beliefs must be such that $U_0^0 = U_1^{0,R}$. Moreover, these beliefs must be derived from equilibrium strategies. Thus, the only $s(\sigma_0)$ that constitutes an equilibrium is obtained by equating (1) and (2), yielding (3).

We now show that, given voters' beliefs, it is indeed optimal for the high-ability politician to choose the policy corresponding to ω. Consider first type $(\theta_H, \omega = 0)$. He prefers a_0 because

$$(1-\phi)\frac{\lambda_I \nu_0}{\lambda_I \nu_0 + (1-\lambda_I)(1-s_R^*(\sigma_0))} > -\phi c$$

for any $s_R^*(\sigma_0)$.

Now consider type $(\theta_H, \omega = 1)$. If $\nu_0 \geq \nu_0^R$, and hence $s_R^*(\sigma_0) = 0$, he prefers a_1 because

$$(1-\phi)\frac{\lambda_I \nu_0}{1-\lambda_I(1-\nu_0)} < \phi(\Delta - c) + 1 - \phi.$$

Turning to the mixed-strategy case where $\nu_0 < \nu_0^R$ and hence $0 < s_R^*(\sigma_0) < 1$, type $(\theta_H, \omega = 1)$ also strictly prefers a_1. Indeed,

$$(1-\phi)\frac{\lambda_I \nu_0}{\lambda_I \nu_0 + (1-\lambda_I)(1-s_R^*(\sigma_0))} < \phi(\Delta - c) + (1-\phi)\frac{\lambda_I}{\lambda_I + (1-\lambda_I)s_R^*(\sigma_0)},$$

because (3) must hold.

Finally, we prove the claimed monotonicity property of the equilibrium strategy by applying the implicit function theorem. Let $F_{0,R} \equiv U_0^0 - U_1^{0,R}$. We have

$$\frac{\mathrm{d}s_R^*(\sigma_0)}{\mathrm{d}\nu_0} = -\frac{\partial F_{0,R}/\partial \nu_0}{\partial F_{0,R}/\partial s}.$$

It is straightforward to see that $\partial F_{0,R}/\partial s > 0$, while $\partial F_{0,R}/\partial \nu_0 > 0$. Hence, $\mathrm{d}s_R^*(\sigma_0)/\mathrm{d}\nu_0 < 0$. ■

Proof of Lemma 2:

The proof is similar to the one for Lemma 1. First, taking the high-ability politician's mechanical play as given, we derive the conditions for pure-strategy equilibria with either $s_R^*(\sigma_1) = 1$ or $s_R^*(\sigma_1) = 0$. We prove the existence of thresholds $\underline{\nu}_1^R$ and $\overline{\nu}_1^R$ as described in the lemma and establish when they are in the range of values that are admissible for ν_1 under Assumption 1. Second, we show that the high-ability politician indeed finds it optimal to follow the equilibrium strategy.

A necessary and sufficient condition for a pure-strategy equilibrium where $s(\sigma_1) = 1$ is that type θ_L prefers a_1 even though voters believe that he will always follow the signal. This condition is obtained by evaluating U_0^1 and U_1^1 at $s(\sigma_1) = 1$, that is

$$\phi[\nu_1 \Delta - c] \geq (1-\phi)\left[1 - \lambda_I \nu_1\right]. \tag{21}$$

A pure-strategy equilibrium with $s(\sigma_1) = 0$ requires that the low type prefer a_0 even though he could fool voters about his type by playing a_1 and succeeding. The associated condition is obtained by evaluating U_0^1 and U_1^1 at $s(\sigma_1) = 0$:

$$\phi[\nu_1\Delta - c] \leq (1-\phi)\left[\frac{\lambda_I(1-\nu_1)}{1-\lambda_I\nu_1} - \nu_1\right]. \tag{22}$$

If neither (21) nor (22) hold, we are in a mixed-strategy equilibrium where the low type randomizes between actions a_0 and a_1.

Since the left-hand side of both (21) and (22) is increasing in ν_1, while the right-hand side of each of those conditions is decreasing in ν_1, the thresholds $\underline{\nu}_1^R$ and $\overline{\nu}_1^R$ exist. Also, $\underline{\nu}_1^R < \overline{\nu}_1^R$ since $\lambda_I(1-\nu_1)/(1-\lambda_I\nu_1) - \nu_1 < 1 - \lambda_I\nu_1$ for any $\lambda_I < 1$. What remains to be shown is when these thresholds are within the set of values consistent with Assumption 1. As $\nu_1 \to 1$, condition (21) is satisfied if $\lambda_I > 1 - \phi(\Delta - c)/(1-\phi)$, while as $\nu_1 \to c/\Delta$, it can never be satisfied. Condition (22), though, cannot be satisfied for $\nu_1 \geq \overline{\nu}_1^R$. As $\nu_1 \to c/\Delta$, it is satisfied if and only if $\lambda_I > c\Delta/((\Delta - c)^2 + c\Delta)$ (see the proof of Lemma 1).

For values between $\underline{\nu}_1^R$ and $\overline{\nu}_1^R$, the equilibrium is in mixed strategies. For a type θ_L politician to be willing to randomize, voters' beliefs must be such that $U_0^1 = U_1^1$. Moreover, these beliefs must be derived from equilibrium strategies. Thus, the only $s(\sigma_1)$ that constitutes an equilibrium is obtained by equating (5) and (6), yielding (7).

We now show that, given the voters' beliefs, it is indeed optimal for the high-ability politician to choose the policy corresponding to ω. Consider first type $(\theta_H, \omega = 0)$. He prefers a_0 because

$$(1-\phi)\frac{\lambda_I(1-\nu_1)}{\lambda_I(1-\nu_1) + (1-\lambda_I)(1-s_R^*(\sigma_1))} > -\phi c,$$

for any $s_R^*(\sigma_1)$.

Now consider type $(\theta_H, \omega = 1)$. If $\nu_1 \geq \overline{\nu}_1^R$, and hence $s_R^*(\sigma_1) = 1$, he prefers a_1 because

$$1 - \phi < \phi(\Delta - c) + (1-\phi)\lambda_I$$

is a necessary condition for the low type to play a pure strategy with $s(\sigma_1) = 1$. In the mixed-strategy case where $\underline{\nu}_1^R < \nu_1 < \overline{\nu}_1^R$ and hence $0 < s_R^*(\sigma_1) < 1$, type $(\theta_H, \omega = 1)$ also strictly prefers a_1 since

$$(1-\phi)\frac{\lambda_I(1-\nu_1)}{\lambda_I(1-\nu_1) + (1-\lambda_I)(1-s_R^*(\sigma_1))} < \phi(\Delta - c) + (1-\phi)\frac{\lambda_I}{\lambda_I + (1-\lambda_I)s_R^*(\sigma_1)}$$

follows from (7). And if $\nu_1 < \underline{\nu}_1^R$ so that $s(\sigma_1) = 0$, type (θ_H, ω_1) has a strict preference for a_1 as well because

$$(1-\phi)\frac{\lambda_I(1-\nu_1)}{1-\lambda_I\nu_1} < \phi(\Delta - c) + 1 - \phi.$$

Finally, we prove the claimed monotonicity property of the equilibrium strategy by applying the implicit function theorem. Let $F_1 \equiv U_0^1 - U_1^1$. We have

$$\frac{\mathrm{d}s_R^*(\sigma_1)}{\mathrm{d}\nu_1} = -\frac{\partial F_1/\partial \nu_1}{\partial F_1/\partial s}.$$

It is straightforward to see that $\partial F_1/\partial s > 0$ and $\partial F_1/\partial \nu_1 < 0$. Hence, $\mathrm{d}s^*(\sigma_1)/\mathrm{d}\nu_1 > 0$. ■

Proof of Lemma 3:

Once again, we will take as given, in a first step, that the high-ability politician plays mechanically as stated in the lemma, and derive the conditions for pure-strategy equilibria with either $s_B^*(\sigma_0) = 0$ or $s_B^*(\sigma_0) = 1$. We show that an equilibrium with $s_B^*(\sigma_0) = 1$ is impossible under Assumption 1; we then prove the existence of a threshold ν_0^B as described in the lemma and establish when it is within the set of values that are admissible for ν_0 under Assumption 1. Second, we show that the high-ability politician indeed finds it optimal to follow the equilibrium strategy.

A necessary and sufficient condition for a pure-strategy equilibrium where $s(\sigma_0) = 0$ is that type θ_L prefers a_0 even though voters believe that he will always follow the signal. From (8) and (9), the condition is

$$\phi[(1-\nu_0)\Delta - c] \leq (1-\phi)\left[\frac{\lambda_I \nu_0}{1-\lambda_I(1-\nu_0)} - \frac{(1-\nu_0)\lambda_I}{\lambda_I + (1-\lambda_I)s_R^*(\sigma_1)}\right]. \tag{23}$$

A pure-strategy equilibrium with $s(\sigma_0) = 1$ would require that the low type prefer a_1 even though he could fool voters about his type by playing a_0. The associated condition is

$$\phi[(1-\nu_0)\Delta - c] \geq (1-\phi)\left[1 - \frac{(1-\nu_0)\lambda_I}{\lambda_I + (1-\lambda_I)s_R^*(\sigma_1)}\right]. \tag{24}$$

If neither (23) nor (24) hold, we are in a mixed-strategy equilibrium where a low type randomizes between actions a_0 and a_1.

As $\nu_0 \to 1$, condition (23) is always satisfied, while as $\nu_0 \to 0$, it can never be satisfied. Together with the monotonicity of both left- and right-hand side of (23) with respect to ν_0, this proves that there exists a unique threshold ν_0^B as stated in the lemma. Condition (24) cannot be satisfied for any $\nu_0 > 1 - c/\Delta$; thus, there is no equilibrium with $s(\sigma_0) = 1$. For the threshold ν_0^B to be greater than $1 - c/\Delta$, the limit of the right-hand side of (19) must be negative as $\nu_0 \to 1 - c/\Delta$ (the left-hand side tends to zero). That is,

$$\lim_{\nu_0 \to 1-c/\Delta} \frac{\lambda_I \nu_0}{1-\lambda_I(1-\nu_0)} - \frac{(1-\nu_0)\lambda_I}{\lambda_I + (1-\lambda_I)s_R^*(\sigma_1)} < 0$$

$$\Leftrightarrow \quad \lambda_I < \frac{(\Delta - c)\Delta(1 - s_R^*(\sigma_1)) + 2c\Delta - \Delta^2}{(\Delta - c)\Delta(1 - s_R^*(\sigma_1)) + c^2}.$$

For values below ν_0^B, the equilibrium is in mixed strategies. For a θ_L type politician to be willing to randomize, voters' beliefs must be such that $U_0^0 = U_1^{0,B}$, with beliefs derived from equilibrium strategies. Thus, the only $s(\sigma_0)$ that constitutes an equilibrium is obtained by equating (8) and (9), yielding equation (10) stated in the lemma.

What remains to be shown is that a high-ability politician chooses to play the policy corresponding to ω, given biased voters' beliefs. Consider first type $(\theta_H, \omega = 0)$. This type prefers a_0 because for any $s_B^*(\sigma_0)$,

$$(1-\phi)\frac{\lambda_I \nu_0}{\lambda_I \nu_0 + (1-\lambda_I)(1-s_B^*(\sigma_0))} > -\phi c.$$

For type $(\theta_H, \omega = 1)$ we need to consider two cases. In the first case, if $\nu_0 \geq \nu_0^B$, in which type θ_L plays the pure strategy $s_B^*(\sigma_0) = 0$, the high-ability politician prefers action a_1 because

$$(1-\phi)\underbrace{\frac{\lambda_I \nu_0}{1-\lambda_I(1-\nu_0)}}_{\leq \lambda_I} < \phi(\Delta - c) + (1-\phi)\underbrace{\frac{\lambda_I}{\lambda_I + (1-\lambda_I)s_R^*(\sigma_1)}}_{\geq \lambda_I}.$$

In the second case, if $\nu_0 < \nu_0^B$, in which a θ_L type plays the mixed strategy $0 < s_B^*(\sigma_0) < 1$, type $(\theta_H, \omega = 1)$ also prefers a_1 since

$$(1-\phi)\frac{\lambda_I \nu_0}{\lambda_I \nu_0 + (1-\lambda_I)(1-s_B^*(\sigma_0))} < \phi(\Delta - c) + (1-\phi)\frac{\lambda_I}{\lambda_I + (1-\lambda_I)s_R^*(\sigma_1)},$$

where the inequality follows from the fact that, in a mixed-strategy equilibrium, the indifference condition given in (10) must hold.

The proof concludes by establishing the claimed monotonicity property of the equilibrium strategy. For this we apply the implicit function theorem. Let $F_{0,B} \equiv U_0^0 - U_1^{0,B}$, then we have

$$\frac{\mathrm{d}s_B^*(\sigma_0)}{\mathrm{d}\nu_0} = -\frac{\partial F_{0,B}/\partial \nu_0}{\partial F_{0,B}/\partial s}.$$

It is straightforward to see that $\partial F_{0,B}/\partial s < 0$, while $\partial F_{0,B}/\partial \nu_0 > 0$. Therefore $\mathrm{d}s_B^*(\sigma_0)/\mathrm{d}\nu_0 > 0$. ■

Proof of Proposition 1:

We start by showing that $\nu_0^R > \nu_0^B$, i.e., the minimum signal precision required to have $s(\sigma_0) = 0$ in equilibrium is greater in the rational than in the biased evaluation regime. This result is immediate when comparing the right-hand side of equations (4) and (11) (cf. Lemma 1 and 3) since, as a result of Assumption 2, $s_R^*(\sigma_1) > 0$ for any ν_1.

Since $\nu_0^R > \nu_0^B$, it follows that $\forall\ \nu_0 \in (\nu_0^B, \nu_0^R)$, we have $s_B^*(\sigma_0) = 0$ by Lemma 1 while $s_R^*(\sigma_0) > 0$ by Lemma 3, and hence the claimed result holds. What remains to be shown is that the claimed result is also true for $\nu_0 \leq \nu_0^B$.

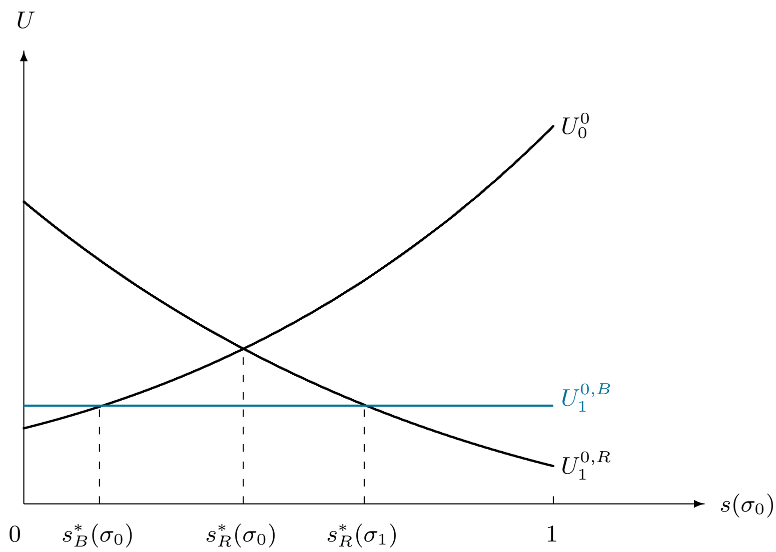

Figure 7: Determination of the low-type's equilibrium strategy

Figure 7 illustrates the logic of the argument. For $\nu_0 \leq \nu_0^B$, the equilibrium is in mixed strategies with both rational and hindsight biased voters. Thus, $s_R^*(\sigma_0)$ is determined by (3) (equating U_0^0 and $U_1^{0,R}$), while $s_B^*(\sigma_0)$ is determined by (10) (equating U_0^0 and $U_1^{0,B}$). U_0^0 is monotone increasing in $s(\sigma_0)$. $U_1^{0,R}$ is monotone decreasing in $s(\sigma_0)$. $U_1^{0,B}$ is constant with respect to $s(\sigma_0)$.

From inspection of $U_1^{0,R}$ and $U_1^{0,B}$, they coincide only at $s(\sigma_0) = s_R^*(\sigma_1)$. Thus, the intersection of U_0^0 and $U_1^{0,B}$, defining $s_B^*(\sigma_0)$, lies to the left of the intersection of U_0^0 and $U_1^{0,R}$, defining $s_R^*(\sigma_0)$, if and only if $s_R^*(\sigma_1) > s_R^*(\sigma_0)$. But this is precisely the subject of Corollary 1; the strict inequality follows from Assumption 2. ■

Proof of Proposition 2:

Given an ex ante signal σ_0, the utility for a θ_L-type politician from playing a_0 does not change from naive hindsight biased (B) to ex post sophisticated (S) evaluation, i.e., $U_0^{0,B} = U_0^{0,S}$. The low type's utility from playing a_1 under sophisticated evaluation (13) can be written as

$$U_1^{0,S} = r\, U_1^{0,B} + (1-r)\, U_1^{0,R}.$$

Thus, it is a convex combination of the payoffs with rational and naive hindsight biased voters. Reliability r, given in (12), satisfies $0 < r < 1$ for any $s(\sigma_0) \in [0,1]$. Moreover, simplifying the expression determining a type-θ_L politician's reputational value under sophisticated evaluation (the second term in (13)) using (12), as

$$r\, \mu(\sigma_1, a_1, \Delta) + (1-r)\, \mu(\sigma_0, a_1, \Delta) = \frac{\lambda_I}{\lambda_I + (1-\lambda_I)[x s_R^*(\sigma_1) + (1-x) s(\sigma_0)]},$$

it is clear that $U_1^{0,S}$ is monotone decreasing in $s(\sigma_0)$. It follows from the proof of Proposition 1 that $s_B^*(\sigma_0) \leq s_S^*(\sigma_0) < s_R^*(\sigma_0)$, where the strict inequality is due to $\nu_0 < \nu_0^R$. ■

Proof of Lemma 4:

We first simplify the expression for $EW_R(s)$. By linearity of $w(\cdot)$, we have $\int_0^\mu w(\mu)\mathrm{d}\lambda_C = \mu w(\mu) = \mu[w(\mu) - w_L]/2 + \int_0^\mu w(\lambda_C)\mathrm{d}\lambda_C$. Thus,

$$\begin{aligned} EW_R(s) &= \int_0^1 w(\lambda_C)\mathrm{d}\lambda_C + \big(\nu_0\lambda_I + (1-\lambda_I)(1-s)\big)\mu_0(s)[w(\mu_0(s)) - w_L]/2 + \\ &\quad + (1-\nu_0)\big(\lambda_I + (1-\lambda_I)s\big)\mu_{1\Delta}(s)[w(\mu_{1\Delta}(s)) - w_L]/2. \end{aligned}$$

Noting that $\int_0^1 w(\lambda_C)\mathrm{d}\lambda_C = (w_H + w_L)/2$ and $\mu(w(\mu) - w_L) = \mu^2(w_H - w_L)$, this expression becomes

$$\begin{aligned} EW_R(s) &= \frac{w_H + w_L}{2} + \frac{w_H - w_L}{2}\Big[\big(\nu_0\lambda_I + (1-\lambda_I)(1-s)\big)\big(\mu_0(s)\big)^2 + \\ &\quad + (1-\nu_0)\big(\lambda_I + (1-\lambda_I)s\big)\big(\mu_{1\Delta}(s)\big)^2\Big]. \end{aligned}$$

Inserting the expressions for the politician's reputation, this simplifies to

$$EW_R(s) = \frac{w_H + w_L}{2} + \frac{w_H - w_L}{2}\left[\frac{\nu_0^2\lambda_I^2}{\nu_0\lambda_I + (1-\lambda_I)(1-s)} + (1-\nu_0)\frac{\lambda_I^2}{\lambda_I + (1-\lambda_I)s}\right],$$

which we can finally rewrite as

$$EW_R(s) = \frac{w_H + w_L}{2} + \frac{(w_H - w_L)\lambda_I}{2}\big[\nu_0\ \mu_0(s) + (1-\nu_0)\mu_{1\Delta}(s)\big]. \tag{25}$$

For the case with hindsight bias, $EW_B(s)$ can be simplified in the same way and becomes

$$\begin{aligned} EW_B(s) &= \frac{w_H + w_L}{2} + \frac{(w_H - w_L)\lambda_I}{2}\Bigg[\nu_0\ \mu_0(s) + \\ &\quad + (1-\nu_0)\mu_{1\Delta}(s_R^*(\sigma_1))\frac{\lambda_I + (1-\lambda_I)(2s_R^*(\sigma_1) - s)}{\lambda_I + (1-\lambda_I)s_R^*(\sigma_1)}\Bigg]. \end{aligned} \tag{26}$$

Comparing the two expectations boils down to

$$\begin{aligned} EW_R(s) &\geq EW_B(s) \\ \Leftrightarrow [\lambda_I + (1-\lambda_I)s_R^*(\sigma_1)]^2 &\geq [\lambda_I + (1-\lambda_I)s][\lambda_I + (1-\lambda_I)(2s_R^*(\sigma_1) - s)]. \end{aligned}$$

Make the following change of variable: $\delta \equiv s - s_R^*(\sigma_1)$ with $\delta \in [-s_R^*(\sigma_1), 1 - s_R^*(\sigma_1)]$. Then, the above inequality is equivalent to

$$[\lambda_I + (1-\lambda_I)s_R^*(\sigma_1)]^2 \geq [\lambda_I + (1-\lambda_I)(s_R^*(\sigma_1) + \delta)][\lambda_I + (1-\lambda_I)(s_R^*(\sigma_1) - \delta)]$$

which is true for any δ and strictly so for $\delta \neq 0$. Hence, $EW_R(s) \geq EW_B(s)$ for any s with strict inequality for $s \neq s_R^*(\sigma_1)$. ■

Proof of Proposition 3:

For $\nu_0 \geq \nu_0^R$, $s_R^*(\sigma_0) = 0$ by Lemma 1, and $s_B^*(\sigma_0) = 0$ by Proposition 1. By Lemma 4, $EW_R(0) > EW_B(0)$, and by continuity, this must also hold when decreasing ν_0 slightly below ν_0^R. ■

Proof of Proposition 4:

This is immediate from the previous discussion given that overall expected welfare is the discounted sum of first- and second-period welfare. ■

Proof of Lemma 5:

Inequality (14) follows directly from the fact that $s_R^*(\sigma_0) > s_B^*(\sigma_0)$ whenever $\nu_0 < \nu_0^R$, a result established in Proposition 1. Pessimistic beliefs ensure that (15) holds for any $s(\sigma_0)$. Inequality (16) is satisfied because $s_R^*(\sigma_0) < s_R^*(\sigma_1)$; see Corollary 1. ■

Proof of Proposition 5:

By applying Lemma 5, we immediately get the required result: since $\mu_0\big(s_R^*(\sigma_0)\big) > \mu_0\big(s_B^*(\sigma_0)\big)$ and $\mu_{1\Delta}\big(s_R^*(\sigma_0)\big) > \mu_{1\Delta}\big(s_R^*(\sigma_1)\big)$ it must be the case that $\mathcal{R}_\theta^B < \mathcal{R}_\theta^R$ for all θ; the probability weights attached to the different posteriors do not matter. ■

B Elimination of alternative equilibria with criterion D1

It is a well-known fact that, because it does not pin down out-of-equilibrium beliefs, the PBE concept is often plagued by multiple equilibria. In our case, there exist equilibria other than those described in Lemma 1 and Lemma 2 satisfying criteria (i) through (iii) of Definition 1; namely, pooling equilibria where all types of politician choose the same policy irrespective of their information. Consider the following sets of strategies and beliefs:

- All types pool on a_0, and voters believe that any politician who plays a_1 is of type θ_L with probability one, i.e., $\mu(\sigma, a_1, y) = 0$;
- all types pool on a_1, and voters believe that any politician who plays a_0 is of type θ_L with probability one, i.e., $\mu(\sigma, a_0, 0) = 0$.

The first of these candidates requires $\frac{1-\phi}{\phi}\lambda_I > \Delta - c$, the second $\frac{1-\phi}{\phi}\lambda_I > c$, to be an equilibrium.

Both of these equilibria can be eliminated using the D1 criterion which puts restrictions on out-of-equilibrium beliefs.[45] We show this for the first of the two candidates (pooling on a_0); the argument can be applied in an analogous way to the other.

Whatever his type, the politician's equilibrium payoff is $(1-\phi)\lambda_I$. Let γ denote a mixed action for the voters, i.e., γ is the probability of voting for the incumbent. Define $D\big((\theta, \Psi_\theta), a_1\big)$ as the set of mixed best responses to action a_1 that makes a politician of type θ and with information Ψ_θ strictly better off playing a_1 than with his equilibrium strategy,

$$D\big((\theta, \Psi_\theta), a_1\big) = \bigcup_{\mu} \{\gamma \in MBR(\mu, a_1) \ : \ (1-\phi)\lambda_I < \phi E(W|\Psi_\theta) + (1-\phi)\gamma\},$$

where $MBR(\mu, a_1)$ is the set of mixed best responses to action a_1 for posterior beliefs μ. Similarly, let $D^0\big((\theta, \Psi_\theta), a_1\big)$ denote the set of responses for which the politician is indifferent. According to the D1 criterion, a type (θ, Ψ_θ) can be deleted for action a_1 if there exists another type $(\theta, \Psi_\theta)'$ (i.e., of different ability or with different information), such that

$$D\big((\theta, \Psi_\theta), a_1\big) \cup D^0\big((\theta, \Psi_\theta), a_1\big) \subset D\big((\theta, \Psi_\theta)', a_1\big),$$

where $\subset$ denotes a strict inclusion.

Let us derive the sets $D\big((\theta, \Psi_\theta), a_1\big)$ for the different types in the most interesting case where $\sigma = \sigma_1$. We have

$$E(W|\Psi_\theta) = \begin{cases} \nu_1\Delta - c & \text{for } \theta_L \\ -c & \text{for } (\theta_H, \omega = 0) \\ \Delta - c & \text{for } (\theta_H, \omega = 1) \end{cases}$$

The voters' best response to a_1 depends on μ. Suppose that the perceived ability of the challenger is λ_C. The voters' best response is "vote for incumbent" if $\mu > \lambda_C$, "vote for challenger" if $\mu < \lambda_C$, and any mixed action $\gamma \in [0, 1]$ if $\mu = \lambda_C$. Thus, any γ is a mixed best response for some belief μ, and

$$\begin{aligned} D\big(\theta_L, a_1\big) &= (\lambda - \tfrac{\phi}{1-\phi}(\nu_1\Delta - c), 1] \\ D\big((\theta_H, \omega = 0), a_1\big) &= (\lambda + \tfrac{\phi}{1-\phi}\, c, 1] \\ D\big((\theta_H, \omega = 1), a_1\big) &= (\lambda - \tfrac{\phi}{1-\phi}(\Delta - c), 1]. \end{aligned}$$

Clearly, if $\nu_1 < 1$, $D\big(\theta_L, a_1\big) \cup D^0\big(\theta_L, a_1\big) \subset D\big((\theta_H, \omega = 1), a_1\big)$, so that type θ_L (and, a fortiori, type $(\theta_H, \omega = 0)$) can be pruned based on criterion D1. That is, voters should believe that a deviation to a_1 is infinitely more likely to come from type $(\theta_H, \omega = 1)$ than from θ_L, in which case they should reelect the incumbent. Anticipating this, the high-ability politician will not stick to his prescribed equilibrium strategy when observing $\omega = 1$, and the equilibrium breaks down.

[45] D1, a refinement developed by Cho and Kreps (1987), is a slightly stronger version of Banks and Sobel's (1987) "divinity" concept.

References

Allen, F. and Gorton, G. (1993), "Churning Bubbles", *Review of Economic Studies*, **60**, 813–836.

Anderson, J. C., Lowe, D. J. and Reckers, P. M. (1993), "Evaluation of Auditor Decisions: Hindsight Bias Effects and the Expectation Gap", *Journal of Economic Psychology*, **14**, 711–737.

Banks, J. S. and Sobel, J. (1987), "Equilibrium Selection in Signaling Games", *Econometrica*, **55**, 647–661.

Barro, R. (1973), "The Control of Politicians: An Economic Model", *Public Choice*, **14**, 19–42.

Bénabou, R. and Tirole, J. (2002), "Self-Confidence and Personal Motivation", *Quarterly Journal of Economics*, **117**, 871–915.

Bénabou, R. and Tirole, J. (2006), "Belief in a Just World and Redistributive Politics", *Quarterly Journal of Economics*, **121**, 699–746.

Besley, T. (2006), *Principled Agents? The Political Economy of Good Government*, Lindahl Lectures on Monetary and Fiscal Policy (Oxford University Press).

Biais, B. and Weber, M. (2006), "Hindsight Bias and Investment Performance", Working paper.

Bryant, F. B. and Brockway, J. H. (1997), "Hindsight Bias in Reaction to the Verdict in the O. J. Simpson Criminal Trial", *Basic and Applied Social Psychology*, **19**, 225–241.

Bukszar, E. and Connolly, T. (1988), "Hindsight Bias and Strategic Choice: Some Problems in Learning from Experience", *Academy of Management Journal*, **31**, 628–641.

Camerer, C., Loewenstein, G. and Weber, M. (1989), "The Curse of Knowledge in Economic Settings: An Experimental Analysis", *Journal of Political Economy*, **97**, 1232–1254.

Canes-Wrone, B., Herron, M. C. and Shotts, K. W. (2001), "Leadership and Pandering: A Theory of Executive Policymaking", *American Journal of Political Science*, **45**, 532–550.

Cho, I.-K. and Kreps, D. M. (1987), "Signaling Games and Stable Equilibria", *Quarterly Journal of Economics*, **102**, 179–221.

Coate, S. and Morris, S. (1995), "On the Form of Transfers to Special Interests", *Journal of Political Economy*, **103**, 1210–1235.

Compte, O. and Postlewaite, A. (2004), "Confidence-Enhanced Performance", *American Economic Review*, **94**, 1536–1557.

Crémer, J. (1995), "Arm's Length Relationships", *Quarterly Journal of Economics*, **110**, 275–295.

DellaVigna, S. and Kaplan, E. (2007), "The Fox News Effect: Media Bias and Voting", *Quarterly Journal of Economics*, **122**, 1187–1234.

Detmer, D. E., Fryback, D. and Gassner, K. (1978), "Heuristics and Biases in Medical Decision-Making", *Journal of Medical Education*, **53**, 682–683.

Dewatripont, M. and Seabright, P. (2006), ""Wasteful" Public Spending and State Aid Control", *Journal of the European Economic Association*, **4**, 513–522.

Downs, G. W. and Rocke, D. M. (1994), "Conflict, Agency, and Gambling for Resurrection: The Principal-Agent Problem Goes to War", *American Journal of Political Science*, **38**, 362–380.

Ferejohn, J. (1986), "Incumbent Performance and Electoral Control", *Public Choice*, **50**, 5–25.

Fischhoff, B. (1975), "Hindsight $\neq$ Foresight: The Effect of Outcome Knowledge on Judgment under Uncertainty.", *Journal of Experimental Psychology: Human Perception and Performance*, **1**, 288–299.

Frey, B. and Eichenberger, R. (1991), "Anomalies in Political Economy", *Public Choice*, **68**, 71–89.

Gowda, M. V. R. (1999), "Heuristics, Biases, and the Regulation of Risk", *Policy Sciences*, **32**.

Harley, E. M. (2007), "Hindsight Bias in Legal Decision Making", *Social Cognition*, **24**, 48–63.

Harrington, J. E. (1993), "Economic Policy, Economic Performance, and Elections", *American Economic Review*, **83**, 27–42.

Hawkins, S. A. and Hastie, R. (1990), "Hindsight: Biased Judgments of Past Events After the Outcomes are Known", *Psychological Bulletin*, **107**, 311–327.

Hoch, S. J. and Loewenstein, G. F. (1989), "Outcome Feedback: Hindsight and Information", *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **15**, 605–619.

Hoffrage, U., Hertwig, R. and Gigerenzer, G. (2000), "Hindsight Bias: A By-Product of Knowledge Updating?", *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **26**, 566–581.

Jacobs, L. R. and Shapiro, R. Y. (2000), *Politicians Don't Pander. Political Manipulation and the Loss of Democratic Responsiveness* (Chicago, IL: University of Chicago Press).

Köszegi, B. (2006), "Ego Utility, Overconfidence and Task Choice", *Journal of the European Economic Association*, **4**, 673–707.

Levy, G. (2004), "Anti-Herding and Strategic Consultation", *European Economic Review*, **48**.

Majumdar, S. and Mukand, S. W. (2004), "Policy Gambles", *American Economic Review*, **94**, 1207–1222.

Maskin, E. and Tirole, J. (2004), "The Politician and the Judge: Accountability in Government", *American Economic Review*, **94**, 1034–1054.

Monroe, A. D. (1998), "Public Opinion and Public Policy, 1980-1993", *Public Opinion Quarterly*, **21**, 6–28.

Mueller, J. E. (1994), *Policy and Opinion in the Gulf War* (Chicago: University of Chicago Press).

Mullainathan, S. (2002), "A Memory-Based Model of Bounded Rationality", *Quarterly Journal of Economics*, **117**, 735–774.

Persson, T. and Tabellini, G. (2000), *Political Economics: Explaining Economic Policy* (MIT Press).

Prat, A. (2005), "The Wrong Kind of Transparency", *American Economic Review*, **95**, 862–877.

Rabin, M. (1998), "Psychology and Economics", *Journal of Economic Literature*, **36**, 11–46.

Rachlinski, J. J. (1998), "A Positive Psychological Theory of Judging in Hindsight", *University of Chicago Law Review*, **65**, 571–625.

Rogoff, K. (1990), "Equilibrium Political Budget Cycles", *American Economic Review*, **80**, 21–36.

Smart, M. and Sturm, D. M. (2006), "Term Limits and Electoral Accountability", Centre for Economic Performance Discussion Paper no. 770.

Viscusi, W. K. (2001), "Jurors, Judges, and the Mistreatment of Risk by the Courts", *Journal of Legal Studies*, **30**, 107–142.