# Submission Number: PET11-11-00061

## Reputation-Concerned Policy Makers and Institution Design

Qiang Fu
*National University of Singapore*

Ming Li
*Concordia University*

## *Abstract*

We study the policy choice of an office-holding politician who is concerned with the public's perception of his capabilities. The politician decides whether to maintain the status quo or to conduct a risky reform. The success of the reform depends critically upon the capability of the politician, which is privately known to the politician. The public observes both his policy choice and the outcome of the reform, and forms a posterior on the true ability of the politician. We show that politicians may engage in socially detrimental reform in order to be perceived as more capable. We investigate the institutional remedy that balances the gains and costs when the policy maker is subject to reputation concerns. Conservative institutions that thwart reform may potentially improve social welfare.

**Submitted:** February 22, 2011.

# Reputation-Concerned Policy Makers and Institution Design*

Qiang Fu†
National University of Singapore
Ming Li‡
Concordia University and CIREQ

First draft: July 2008
This version: July 2010

**Abstract**

We study the policy choice of an office-holding politician who is concerned with the public's perception of his capabilities. The politician decides whether to maintain the status quo or to conduct a risky reform. The success of the reform depends critically upon the capability of the politician, which is privately known to the politician. The public observes both his policy choice and the outcome of the reform, and forms a posterior on the true ability of the politician. We show that politicians may engage in socially detrimental reform in order to be perceived as more capable. We investigate the institutional remedy that balances the gains and costs when the policy maker is subject to reputation concerns. Conservative institutions that thwart reform may potentially improve social welfare.

JEL Nos: C72, D72, D82
Keywords: Reform, Reputation, Ability, Conservatism

*We are grateful to Parimal Bag, Michael Baye, Oliver Board, Adam Brandenburger, Kim-Sau Chung, Rick Harbaugh, Navin Kartik, Kyungmin Kim, Tilman Klumpp, Jingfeng Lu, Marco Ottaviani, Ivan Png, Larry Samuelson, Joel Sobel, and Bernard Yeung for their helpful comments. We gratefully acknowledge valuable input from the seminar and conference audiences at the Midwest Economic Theory Meeting–Fall 2008, the Canadian Economic Association Meeting–2009, the Far Eastern Meeting of the Econometric Society–2009, the Montreal Theory Reading Group, the Hong Kong University of Science and Technology, Stony Brook Workshop on Political Economy–2010, and Theory Lunch at UW-Madison. Part of the research was completed while Fu was a CIREQ visitor, and he is grateful for the hospitality.

†Department of Strategy and Policy, National University of Singapore, 15 Kent Ridge Drive, SINGAPORE, 119245, Tel: (65)65163775, Email: bizfq@nus.edu.sg.

‡Department of Economics, Concordia University, 1455 Boulevard de Maisonneuve Ouest, Montreal, QC, H3G 1M8, CANADA, Tel: +1(514)8482424 Ext. 3922, Email: mingli@alcor.concordia.ca.

# 1 Introduction

> She (Emma) was not much deceived as to her own skill either as an artist or a musician, but she was not unwilling to have others deceived, or sorry to know her reputation for accomplishment often higher than it deserved.
>
> *Emma*, vol. 1, ch. 6, by Jane Austen, English Author

> Love of fame brings about eccentricity, and being eccentric brings danger to oneself; therefore the sages exhorted against the love of fame.
>
> *Xing xin za yan*, Li Bangxian, Chinese Poet

We are often concerned about the inferences that people make about us based on our actions and the consequences of these actions. These inferences shape our reputations and often determine our prospects of success, professional or otherwise. Reputation concerns loom large, perhaps more conspicuously, in the public sector or in non-profit organizations, where formal contracts based on explicit performance-based incentives are usually rare.

This paper identifies one particular context in which reputational concerns affect economic agents' behaviour. When individuals responsible for public policy are motivated by the concerns for their individual reputations, they may embark on innovative but risky initiatives ("reforms") to convince the public of their capability. Such initiatives, however, can make the public worse off. To prevent the potentially negative consequences of such risky behaviour, it may be necessary to enact "conservative" political and social institutions that restrict policy makers' discretion. Such institutional conservatism may have to reject valuable reform proposals that, if implemented, would benefit the society.

There are many examples of both the strength and prevalence of reputation concerns. The reputation of a technocrat's professional competence often determines his ability to either climb the hierarchy, or to resume his career in the private sector after his term of service is over.[1] Career politicians provide even more salient examples. The prospect of a politician being re-elected depends to a large extent on the public's perception of his capabilities. For example, in the aftermath of the economic turmoil, Gordon Brown was said to have lost his "reputation for economic competence" "through a combination of appallingly bad luck and even worse misjudgment,"[2] which eventually cost him his premiership. Alternatively, a politician in office may have strong concerns about how the public evaluates his legacy when

[1] Bureaucrats in the Securities and Exchange Commission, for instance, often seek lucrative job offers from private financial firms after they leave office.

[2] Source: Fraser Neslon, "Brown's Reputation for Economic Competence Has Gone. The Tories Should Seize the Chance." http://www.spectator.co.uk, January 23rd, 2008.

he steps down. Under various circumstances, reputation concerns are an important part of the informal incentives that politicians and other economic agents face.[3]

In this paper, we examine policy makers' incentives to implement risky policies ("reforms"), which are likely to affect the public's perception of their talent. The success of such reforms depends on an individual policy maker's capabilities. We show that he would carry them out even if he knows that he has a poor chance of success. As Tereza Capelos (2005) states, "political actors often engage in controversial activities", even those that "challenge their reputations." She points out that politicians risk losing their support "after showing inexperience or wrong judgment." Our analysis predicts that such risky actions can be interpreted as rational attempts to enhance one's reputation. Reputational concerns cause excessive and inefficient risk-taking and incur costs to social welfare. In addition to delineating policy makers' behaviours, we investigate the institutional remedies necessary to balance the gains and costs when policy makers are subject to these incentives.

Throughout this paper, the decision maker is generically referred to as a "politician," who chooses between adopting a risky policy option, referred to as "reform," and maintaining the status quo.[4] His policy performance is independent of his capability when the status quo is maintained. However, his inherent capabilities come to the forefront when the status quo is abandoned. The performance of reform depends on not only the intrinsic value of the available proposal, but also on how well he implements such reforms. For instance, if the U.S. President pushes through a fiscal stimulus plan which may help rescue an economy in recession, its ultimate success depends largely on how funds are allocated to optimize its effectiveness.[5] A new policy increases uncertainty, and its success depends on his ability to gather information and take appropriate action under each contingency.[6] A capable politician is thus better at implementing reform, and therefore more likely to be successful.[7]

[3]For instance, Frederick Sheehan (2009) commented that Alan Greenspan deliberately built up his own reputation of competence in designing monetary policy, and went to great lengths to protect it.

[4]Our analysis examines this issue in a variety of environments, including a judge who has to decide whether to exercise his power to strike down a law, a prosecutor who has to decide whether to file charges against a crime suspect, a CEO who has to decide whether to implement an expansion plan, and a a doctoral candidate who must decide whether to pursue a cutting-edge research project.

[5]In another example, although the acquisition of Compaq by Hewlett-Packard has shown its merit over the years, it is widely believed that the initial fiasco was due to the flawed management of the merger by its CEO then, Carly Fiorina.

[6]We do not model how politicians gather information. However, a politician's capability to elicit information from various sources is widely viewed as an important part of leadership. The US presidential historian, Erwin C. Hargrove (1966, pp 70-73 and pp 114-116), paints two completely different pictures of Franklin D. Roosevelt and Herbert Hoover with respect to information gathering. Roosevelt brought together experts who held a great variety of views and balanced them off against each other, while Hoover did not enjoy obtaining critical advice from anyone.

[7]This assumption can be related to the concept of "state capacity" proposed by Theda Skocpol (1985). She

A politician's capability level can be either high or low, and it is known only to himself. The public forms an assessment of this capability based on observations of both the policy choices of the politician and the resultant performance. The politician makes his policy choice to maximize the public's perception of his competence.

Our analysis proceeds in two layers. We first depict the equilibrium behaviour of the politician, which is summarized as follows.

- **No full separation.** There exists a unique semi-separating equilibrium. A politician with a high level of capability (henceforth, the high-type politician) is always "eager" to reveal more information by undertaking reform: he carries out reforms with probability one whenever a sufficiently valuable proposal is available. The low-capability politician (henceforth, the low-type politician) mimics the behaviour of his high-type counterpart, even though there is a higher probability of failure. Reputational concerns "force" him to take risks, because not acting would cause him to suffer from a more unfavourable public assessment.

- **Pressure to prove oneself.** The low-type politician reforms less frequently when the public holds a more favourable prior view of his capability. Because of this effect, without knowing the politician's true type, reform can be predicted to occur less often when the initial assessment of the politician is more favourable. We interpret these results as an indication of a *pressure to prove oneself* phenomenon. It is commonly observed in the intellectual, political, and social aspects of our lives. We discuss it more extensively in Section 3 along with illustrative examples.

- **Tough act to follow.** The higher the capability differential between the high-type and the low-type politician, the less likely it will be that the low-type politician undertakes reform. As successful mimicry becomes more difficult, the low-type politician reforms less often to avoid failure.

Based on the equilibrium results, we evaluate the ramifications of reputation incentives on social welfare. We consider the design of the optimal welfare-maximizing institutions (e.g. constitution) or bureaucratic rules that restrict the discretion of the politician. The results are summarized as follows.

- **Thwarted good reforms.** Assume that a "legislature" e.g., parliament, supreme court, advisory committee, or board of directors, regulates and monitors the policy choice of the politician. The legislature abides by a "constitution" that is embodied

argued that ambitious reform attempts often fail because bureaucrats usually lack the required competence to administer their reform.

through a threshold rule – it prohibits reform unless the intrinsic value of the reform proposal exceeds a threshold. A higher, or more conservative, threshold discourages an incapable politician from undertaking detrimental reform, but it also prevents a capable politician from undertaking beneficial reform. The social optimum requires a proper level of "institutional conservatism" such that the optimal threshold rule must thwart otherwise beneficial reform. Our analysis lends support to the institutions or bureaucratic rules present in various organizations that restrict the ability of politicians or bureaucrats to carry out risky activities at their discretion.[8] It also provides an alternative rationale for the often observed organizational resistance to policy reform and the widely discussed bias towards the status quo. As pointed out by Raquel Fernandez and Dani Rodrik (1991), "one of the fundamental questions in political economy" has been why governments often fail to carry out efficiency-enhancing reform.

- **Opportunities hurt and "optimism" requires more caution**. In an environment in which good reform proposals are more likely to emerge, social welfare could turn out to fall. Low-type politicians are "forced" to reform more often, as the choice to forego reform will be more likely to be attributed by the public to the politician's lack of ability, instead of the lack of opportunities, i.e., the reform proposal is of low value. The joint effect may be that a more favourable environment paradoxically leads to decreases in social welfare. To remedy this problem, more conservative institutional rules may be necessary.

In the rest of this section, we discuss the link between our paper and the relevant literature. In Section 2, we set up the model. In Section 3, we characterize equilibria of the model and present comparative statics of relevant environmental factors. In Section 4, we discuss the welfare implications of our equilibrium results and consider the issue of institutional design. In Section 5, we investigate robustness of our findings and compare our paper with closely related work. We conclude in Section 6. All proofs are collated in the Appendix.

## Relationship to the Literature

The notion of career or reputation concerns is featured prominently in the pathbreaking work of Bengt Holmström (1982, 1999). Since then, an enormous amount of scholarly effort has been devoted to exploring the incentive effects of reputation or career concerns in a wide array of environments, including corporate decision making (e.g., Bengt Holmström and Joan E. Ricart i Costa 1986, Jeffery Zwiebel 1995, and Adam Brandenburger and Ben

[8]For instance, a key issue in the debate between judicial restraint and judicial activism is whether judges should be encouraged to refrain from exercising their power to strike down existing laws.

Polak 1996), economic agents' effort supply (e.g., Holmström 1999 and Alberto Alesina and Guido Tabellini 2007), and experts' strategic advising activities (e.g., Stephen Morris 2001 and Marco Ottaviani and Peter Norman Sørensen 2006). The literature reveals in various contexts that concerns regarding public or market perceptions distort economic agents' decision making. Such incentives lead economic agents' to ignore their own useful information and instead, to strategically manipulate the belief of the public or the market.[9]

Our paper explores (1) the politician's incentives to conduct reform, so as to signal his competence; and (2) the welfare-maximizing institutional rule that restricts the politician's discretion when he is subject to reputation concerns. Hence, it belongs to the strand of career concerns literature that focuses on agents' incentives to undertake risky projects. The setup of our paper is a variation of the example introduced in Section 3.2 of Holmström's (1999) seminal paper. The common feature is that the politician's (decision maker's) talent is only relevant when the reform (risky project) is undertaken. Hence, more information can be revealed when the risky activity is carried out.[10] Two features distinguish our setup from Holmström's (1999): first, we assume the politician's talent is his private information, while he assumes symmetric information and symmetric information updating; second, in our model, the probability of success for each type is common knowledge, while in Holmström's (1999), it is the private information of the agent. As a consequence, in our model, the choice to undertake reform can signal the type of the politician, which is not possible in Holmström's (1999) setup.

Holmström and Ricart i Costa (1986) study managers' incentives to make new investment when the manager is subject to career concerns. Benjamin E. Hermalin (1993) shows that a risk-averse agent with career concerns may have the incentive to choose a more risky project. However, in his model, a risky project is a *worse* indicator of the agent's talent, while in ours, reform reveals *more* information. Gary Biglaiser and Claudio Mezzetti (1997) study politicians' incentives to implement major new projects. Voters evaluate the incumbent's ability based on his performance in the project. They show that the incumbents' project choices can be either too radical or too conservative, depending on the bias of the median voters. Similar to Holmström's (1999), these studies mainly focus on symmetric-information settings, where the type of the manager is unknown to all players. Biglaiser and Mezzetti (1997) also briefly discuss an extension of asymmetric information. They demonstrate the impossibility of full separation but do not fully characterize all the equilibria in that case. Within this strand of literature, our paper is closely linked to that of Jeffery Zwiebel (1995). He also explores how

[9]For instance, Brandenburger and Polak (1996), David S. Scharfstein and Jeremy C. Stein (1990), and Ottaviani and Sørensen (2006) all share this feature.

[10]The inclusion of a "status quo" option that does not reveal the right action to take for the risky option is also present in Amal Sanyal and Kunal Sengupta (2006). They study a game of strategic communication in which the expert is career-concerned in the sense of Ottaviani and Sørensen (2006).

reputation concerns moderate a manager's incentive to undertake risky innovation. In his setting, the manager's innovative action is unobservable and does not signal the manager's type. As a result, he shows that managers may resist beneficial innovation; while we arrive at the opposite conclusion.[11]

Our study includes flavours from both the literature of signalling and that of career concerns, which places it in the company of a handful of other studies. They include the notable examples of Canice Prendergast and Lars Stole (1996), Gilat Levy (2007), and Wei Li (2007).[12] In a recent paper, Kim-Sau Chung and Péter Esö (2008) build a model in which a worker chooses a task to both signal his capabilities to potential employers and learn about his capabilities himself, as he has only imperfect knowledge of it. They assume that the more difficult task is a worse (less informative) device for assessing the capability of a worker; meanwhile in our setting, undertaking the more difficult task (reform) allows for more information transmission.

Sumon Majumdar and Sharun W. Mukand (2004) and Guido Suurmond, Otto H. Swank, and Bauke Visser (2004) both consider the incentives of agents in the public sector to undertake risky projects, which signal their types. Majumdar and Mukand (2004) study the dynamic incentives of a government within an election cycle and its policy persistence. The government can be either too radical or too conservative in equilibrium. Suurmond, Swank, and Visser (2004) contend that the presence of career concerns can be socially beneficial, as it can encourage a smart agent to expend more effort in gathering information. Ying Chen (2010), in a simultaneous and independent paper, analyzes the choice of an agent between a risky project and a safe project. Chen (2010) mainly focuses on the impact of information structure on the agent's project choice under career concerns. She shows that the agent takes excessive/inadequate risks when he does/does not know his own type. In contrast, we focus on a setting in which the politician knows his type. In addition to identifying the problem of excessive risk taking, we focus more on the roles of various environmental factors in determining the politician's behaviour and the design of optimal institution that remedies this problem. Besides the difference in focuses, our study exhibits different modeling characteristics from Majumdar and Mukand (2004), Suurmond, Swank, and Visser (2004) and Chen (2010). The modeling difference and the roles played by the unique flavours of our model will be discussed in Sections 2 and 5.

Our analysis of optimal institution design in the presence of reputation concerns is con-

[11]Robert A. J. Dur (2001) and Peter Howitt and Ronald Wintrobe (1995) also explore scenarios in which there is too little change in policy.

[12]Prendergast and Stole (1996) argue that career concerns induce young investors to overreact to new information they receive, so as to signal that they are fast learners. Wei Li (2007) makes a similar point in the case of experts providing advice to decision makers. We discuss Levy's work at the end of the literature discussion.

ceptually related to that in a small number of other papers, which study the ramifications of various institutional elements in career-concerns models. Andrea Prat (2005) argues that transparency in an organization may hurt as the agent may take revealed action to influence the principal's posterior instead of seeking the best interests of the organization.[13] Gilat Levy (2007) shows that in a committee of voters with career concerns, radical actions are more likely to be accepted when the voting process is transparent. Felix Bierbrauer and Lydia Mechtenberg (2008) analyze the welfare effect of early elections when the political leaders have career concerns. To our knowledge, our paper might be one of the first to explicitly investigate an institutional remedy for inefficient risk taking when the decision maker has reputation concerns. Our result that restrictions on changes to the status quo could be welfare-improving complements other rationales of institutional conservatism, for example, those offered by Li, Hao (2001) and Young K. Kwon (2005). Our analysis espouses the merit of institutional barriers (bureaucracy) that limit policy makers' discretionary power. The paper echoes the conclusion of Jean Tirole (1986) in this respect.

# 2 Setup

A politician makes a policy choice between two alternatives: maintaining the status quo or initiating a reform. If the politician retains the status quo, the outcome of this policy, $y$, is deterministic, which we normalize to 0. In contrast, if the politician chooses to undertake the reform, uncertainty will arise and the politician must take an action to address it. The outcome of a reform is given by the widely adopted quadratic loss function

$$y = \theta - (a - \omega)^{2}. \tag{1}$$

where $\theta$ measures the intrinsic value of the available reform proposal, $\omega$ is the true state of the world, and $a$ is the action taken by the politician in response to his assessment of $\omega$.

## 2.1 Information Structure

The intrinsic value of the available reform proposal, $\theta$, is continuously distributed on $[-\theta_1, \theta_2]$ with a distribution function $F$ and density function $f$, where $-\theta_1 < 0 < \theta_2$ and $\theta_1, \theta_2 \in (1, 2)$. The distribution of $\theta$ is common knowledge. The realization of $\theta$ is observed by the politician before he decides whether or not to adopt the reform proposal, while it is unobservable or *ex ante* unverifiable to the general public.

The ultimate consequences of the reform depend not only on the intrinsic value of the reform proposal, but also the quality of the politician's implementation, i.e., how well he

[13] In a simple and straightforward extension of our basic model, we can also demonstrate that transparency leads to harmful outcomes under certain circumstances.

addresses the uncertainty that arises with reform. The uncertainty is embodied by the state of the world, $\omega$, which may take either of two values from $\Omega = \{-1, 1\}$, each with a probability 1/2. The state $\omega$ is realized only after a reform has been initiated. The politician has to choose his action $a$ from $A = \{-1, 1\}$ to implement the reform. We say that the reform is a *success* if his action matches the state of the world, while it is a *failure* if it does not. Neither the politician nor the public observes the true state. However, the politician can receive a signal $\sigma \in \{-1, 1\}$ about $\omega$. Upon receiving $\sigma$ (either informative or uninformative), the politician takes an action.[14]

The precision of the signal depends on the talent of the politician. The talent of the politician, $t$, is drawn from the set $\{L, H\}$. A high-talent politician ($H$) receives an informative signal, which matches the true state with a probability $q = \Pr(\sigma = \omega) > 3/4$. In contrast, a low-talent politician's signal is completely uninformative. It should be noted that the assumption $q > 3/4$ does not affect our analysis. However, without this assumption, no reform can be socially beneficial. The talent of the politician is his private information. Let $\alpha$ be the probability of $t = H$, which is commonly known. It is the public's prior about the politician's talent, which can also be viewed as the proportion of high-capability politicians in the "population."[15] We assume that the proportion of "good" politicians in the population is small, i.e. $\alpha < \frac{1}{2}$.[16]

The public observes the politician's policy choice (status quo or reform) and the final outcome $y$.[17] By the assumptions of $-\theta_1 < 0 < \theta_2$ and a quadratic output function, $\theta$ can be ex post inferred from $y$ if and only if the reform is carried out. It allows the public to learn whether the reform succeeded or failed. However, the value of $\theta$ is unknown to the public if no reform is undertaken. In Section 5, we discuss an extension to the basic setting where it can be learned without actual reform. We demonstrate that such "transparency" leads to welfare loss and an efficient institution should never allow it.

[14]The distinction between policies (status quo or reform) and actions is important in our model. Policies are macro-level or "strategic" decisions such as whether to reform financial regulations or whether to start a war. In contrast, actions are micro-level or "tactical" decisions such as which instrument of regulation to introduce in overhauling the financial system or how many troops to deploy in the war. The true nature of the problem ($\omega$) determines which action is ex post suitable for implementing the reform.

[15]There is literature that analyzes the composition of politicians as a group, which is complementary to our research, in that it offers an explanation for why politicians may consist of a significant proportion of low-ability individuals. Francesco Caselli and Massimo Morelli (2004), Matthias Messner and Mattias K. Polborn (2004), and Andrea Mattozzi and Antonio Merlo (2007, 2008) have offered various explanations for why political processes tend to select low-ability individuals to be politicians.

[16]This regularity assumption is only required so that in the extreme case where the high type's signal is perfectly informative, the low type still has an incentive to undertake reform and mimic the high type (see the proof of Part 1 of Proposition 1.)

[17]In our setup, whether or not the public observe the action is inconsequential. Once the politician chooses reform, the belief of the public is determined only by whether the outcome is a "failure" or "success."

## 2.2 Payoff

Based on the observation of the politician's policy choice and the subsequent performance, the public's forms a posterior on the type of the politician, which is written as

$$\mu^i(y) \equiv \Pr(t = H | \, y, i)$$

by Bayes' rule, where $i = 0$ indicates the status quo and $i = 1$ indicates reform. Borrowing from much of the career-concerns literature, we assume that the politician's payoff depends purely on his reputation $\mu^i(y)$.

## 2.3 Action Space

We assume that the politician has only limited discretion. He is subject to an institutional constraint and is authorized to undertake a reform only if the intrinsic value of the available reform proposal exceeds a cutoff $\hat{\theta}$. We implicitly assume that the politician's policy choice is subject to the regulation or monitoring of a legislature, e.g., parliament, supreme court, advisory committee, or board of directors. The legislature cannot verify the type of the politician but it can verify the value of the reform proposal, and it abides by certain institutional rules that constrain the politician's authority or discretion. The rules can be understood as a constitution, or as widely observed, an organizational bureaucracy (see Tirole 1986), which prevents him from choosing apparently harmful policies. Such institutional restrictions are prevalent in political and public life. For instance, the US President must obtain congressional approval for his policy choices. Military commanders have to honour "rules of engagement" in the use of force. An administrator of the Environmental Protection Agency has limited authority and resources to regulate polluting industries. Finally, judges are often pressured to refrain from exercising their power to strike down existing laws.

Analogous to Tirole (1986), we implicitly assume that the politician must provide a verifiable report on the value ($\theta$) of his reform proposal to the legislature when he pushes forward a reform, although such information is neither verifiable nor accessible ex ante to the general public. To abide by the "constitution," the legislature would not approve any reform proposal with a value below $\hat{\theta}$. For the moment, we assume that $\hat{\theta}$ is fixed and focus on the equilibrium behaviour of the politician. We dedicate Section 4 to an in-depth analysis on the welfare-maximizing rule $\hat{\theta}^*$, which endogenizes the cutoff.

## 2.4 Timeline

The timeline of the model is as follows.

1. Nature chooses the quality of the reform proposal, i.e., the value of $\theta$.

2. The politician observes $\theta$ and decides whether to adopt the reform proposal. He further chooses $a$ if he decides to undertake the reform.

3. The public updates their belief after observing both the politician's policy choice and performance.

### 2.5 Remark on Model Setup

While the setup of our model differs from the existing literature in a number of ways, we would like to stress one essential feature of our model. In contrast to many existing career-concerns models with risky experimentation (e.g., Majumdar and Mukand 2004, Suurmond, Swank, and Visser 2004), where the outcome of a reform is measured as a binary indicator (e.g., success or failure) alone, we assess its performance as a continuous variable. The performance of a reform proposal depends both on the quality of the reform proposal ($\theta$) and the quality with which it is implemented ($|a - \omega|$), with both subject to random perturbation. This setup enriches our analysis in two aspects. First, it enables an analysis of institutional design. A more sophisticated trade-off is involved in determining the proper level of institutional conservatism. Second, a comparative static analysis may be performed on the probability distribution of the value of reform, which sheds further light on equilibrium behaviour and the design of welfare-maximizing institutions.

## 3 Equilibrium Analysis

In this part, we first study the benchmark of the first best situation in which the public's expected payoff from the politician's policy choice is maximized. We then derive the equilibrium of the game and conduct comparative analysis.

### 3.1 First Best Benchmark

Let us define

$$q_t = \begin{cases} q & \text{for } t = H; \\ \frac{1}{2} & \text{for } t = L. \end{cases}$$

When a high-type politician chooses to reform, he would maximize his probability of success by following his signal, i.e., choosing $a = \sigma$. A low-type politician's signal is uninformative and the two states are equally likely. His choice of $a$ is *ex ante* irrelevant. Hence, the expected outcome of the reform is given by

$$\begin{aligned} E(y) &= \theta - E_{\omega \in \{-1,1\}}(a - \omega)^2 \\ &= \theta - 4(1 - q_t) \end{aligned}$$

In the first-best situation, a politician would undertake reform if and only if the expected outcome $E(y)$ is non-negative. A low-type politician should never reform regardless of $\theta$ as the expected loss from wrong actions always exceeds the benefit of reform, that is,

$$E(y) = \frac{1}{2}\theta + \frac{1}{2}(\theta - 4) = \theta - 2 < 0,$$

because $\theta \leq \theta_2 < 2$. The expected outcome for a high-type politician is given by

$$E(y) = \theta - 4(1 - q).$$

The high type should undertake reform if and only if the value of reform is sufficiently high, i.e., $\theta \geq 4(1 - q)$.

## 3.2 Equilibrium

We adopt the solution concept of Divine Equilibrium, first introduced by Jeffrey S. Banks and Joel Sobel (1987). The equilibrium requires (1) the politician and the public to form Bayesian beliefs, (2) the politician to choose the action that maximizes his expected reputation if he undertakes reform, (3) the politician to choose reform or the status quo so as to maximize his expected reputation; and (4) the out-of-equilibrium belief to satisfy the "divinity" criterion. The "divinity" criterion imposes mild and sensible restrictions on out-of-equilibrium beliefs, which refine the Perfect Bayesian Equilibrium. The basic idea is that if an off-equilibrium action is "more likely" to benefit a certain type of politician, then the public must assign a higher likelihood to that type of politician taking that particular action. However, the "divinity" criterion is weaker than the popularly adopted D1 criterion of Banks and Sobel (1987). Further details are provided in the Appendix.

### 3.2.1 Equilibrium Characterization

First, we consider the politician's strategy. Let $\rho_t(\theta)$ be the probability with which a type-$t$ politician chooses reform when its value is $\theta$. As implied by the institutional rule, $\rho_t(\theta) = 0$ for $\theta \in [-\theta_1, \hat{\theta})$, for $t \in \{L, H\}$. When the politician maintains the status quo, his reputation among the public is

$$\mu^0 = \frac{\alpha F(\hat{\theta}) + \alpha \int_{\hat{\theta}}^{\theta_2} [1 - \rho_H(\theta)] f(\theta) d\theta}{\left[ F(\hat{\theta}) + \alpha \int_{\hat{\theta}}^{\theta_2} [1 - \rho_H(\theta)] f(\theta) d\theta + (1 - \alpha) \int_{\hat{\theta}}^{\theta_2} [1 - \rho_L(\theta)] f(\theta) d\theta \right]}. \tag{2}$$

Note that, as long as reform is undertaken, the public can *ex post* perfectly infer the value of $\theta$ from the outcome $y$. When the politician implements a reform of value $\theta$, his reputation will become

$$\mu^s = \frac{\alpha q \rho_H(\theta) f(\theta)}{\alpha q \rho_H(\theta) f(\theta) + (1 - \alpha) \frac{1}{2} \rho_L(\theta) f(\theta)}$$

if the reform succeeds, and

$$\mu^f = \frac{\alpha(1-q)\rho_H(\theta)f(\theta)}{\alpha(1-q)\rho_H(\theta)f(\theta) + (1-\alpha)\frac{1}{2}\rho_L(\theta)f(\theta)}$$

if the reform fails. If a type-$t$ politician implements a reform with value $\theta$, he receives an expected payoff

$$\mu_t = q_t\mu^s + (1-q_t)\mu^f.$$

The following proposition characterizes the full set of equilibria.

**Proposition 1.** *1. For each given cutoff $\hat{\theta} \in [-\theta_1, \theta_2]$, there exists a unique equilibrium of the game. In equilibrium, the high-type politician undertakes reform with probability $\rho_H^*(\theta) = 1$ whenever he receives a proposal of value $\theta \in [\hat{\theta}, \theta_2]$ and $\rho_H^*(\theta) = 0$ otherwise, while the low-type politician undertakes reform with a probability $\rho_L^*(\theta) = \rho^* \in (0,1)$ when $\theta \in [\hat{\theta}, \theta_2]$ and $\rho_L^*(\theta) = 0$ otherwise.*

*2. The equilibrium probability $\rho^*$ solves*

$$\frac{1}{1+\lambda(\alpha)A} = \frac{1}{2}\cdot\frac{1}{1+\lambda(\alpha)B} + \frac{1}{2}\cdot\frac{1}{1+\lambda(\alpha)C}, \tag{3}$$

*where*

$$\lambda(\alpha) = \frac{1-\alpha}{\alpha},\ A = 1+(1-\rho)\kappa(\hat{\theta}),\ \kappa(\hat{\theta}) = \frac{1-F(\hat{\theta})}{F(\hat{\theta})},\ B = \frac{\frac{1}{2}\rho}{q},\ C = \frac{\frac{1}{2}\rho}{1-q}.$$

Proposition 1 states that there can be no full separation of the two types. The high type is always "eager" to undertake reform: he does so whenever a sufficient valuable proposal is received, i.e. $\theta \geq \hat{\theta}$. Also, whenever $\theta \geq \hat{\theta}$, the low type mimics his high-type counterpart and undertakes reform with a positive probability, $\rho^*$. It should be noted that the equilibrium predicted for each particular $\hat{\theta}$ remains one of the Perfect Bayesian Equilibria in an alternative model without institutional constraints. The presence of institutional constraints does not alter the main property of the equilibrium, i.e. no full separation. However, this setup economizes on our presentation and allows us to concentrate on institutional design.

The policy choice for the low-type politician involves subtle trade-offs. Full separation is impossible in equilibrium. low-type politician cannot completely abstain from undertaking the reform in equilibrium: the absence of reform would be seen as a sign of incompetence while the enactment of reform would be interpreted as a sure sign of competence. Full pooling, however, is not possible either. The low-type politician has a lower chance of success with reform than the high-type politician. Further, if no reform is taken, the public cannot perfectly infer the politician's true type: the realization of $\theta$ is not observable to the public in this case; while the value of $\theta$ can fall below $\hat{\theta}$, which would prevent both types from undertaking reform. The low-type politician thus randomizes in equilibrium.

The nature of the strategic concerns is better revealed when we compare the equilibrium behaviours under different $\hat{\theta}$. Denote by $E\mu_t(\hat{\theta})$ the ex ante expected payoff of a type-$t$ politician in an equilibrium with a given $\hat{\theta}$. Our analysis leads to the following.

**Proposition 2.** *1. The equilibrium probability of reform by the low type, $\rho^*$, strictly decreases with $\hat{\theta}$.*

*2. The low-type politician always prefers a higher cutoff $\hat{\theta}$, while the high-type politician always prefers a lower $\hat{\theta}$. That is, $\frac{dE\mu_H(\hat{\theta})}{d\hat{\theta}} < 0$, and $\frac{dE\mu_L(\hat{\theta})}{d\hat{\theta}} > 0$.*

The politician faces a more stringent standard for taking reform when a higher $\hat{\theta}$ is in place. Overall, high type reforms less when $\hat{\theta}$ increases. Thus the public is more likely to interpret a no-reform outcome as the result of a lack of reform opportunities ($\theta < \hat{\theta}$), rather than the poor capabilities of politicians. Hence, the low-type politician obtains higher reputation from maintaining the status quo, which causes him to reform with a smaller probability.

Part 2 of Proposition 2 describes the two types' utility ranking under different cutoff rules $\hat{\theta}$. The high-type politician prefers a lower cutoff, which allows him to reform more. His capability is revealed with a higher probability. The low-type politician, however, prefers a higher cutoff, and hence less reform, for two reasons. First, the lower frequency of reform allows the low-type politician to pool with his high-type counterpart more often and to reveal less information. Second, when the low-type politician reforms less often, i.e. $\frac{d\rho^*}{d\hat{\theta}} < 0$, the public would believe that a reform is increasingly likely to be implemented by the high-type politician, which mitigates the damage to the low type when his reform fails.

### 3.2.2 Comparative Statics

We now examine how the politician's equilibrium behaviour varies with environment parameters. In equilibrium, the high-type politician reforms with probability one whenever $\theta$ exceeds $\hat{\theta}$, while the low-type politician reforms with a probability $\rho^*$. Hence, in this equilibrium, reform occurs with a probability

$$\bar{\rho} = [1 - F(\hat{\theta})][\alpha + (1 - \alpha)\rho^*]. \tag{4}$$

The main results are summarized in the following proposition.

**Proposition 3.** *Consider the equilibrium under a threshold rule $\hat{\theta}$.*

*1. The probability of reform by the low-type politician, $\rho^*$, is strictly decreasing in $\alpha$, the public's prior. The overall likelihood of reform, $\bar{\rho}$, also strictly decreases with $\alpha$.*

2. *The probability of reform by the low type, $\rho^*$, is strictly decreasing in $q$. The overall likelihood of reform, $\bar{\rho}$, also strictly decreases with $q$.*

3. *Let $\rho$ and $\rho'$ denote, respectively, the equilibrium probabilities of the low type undertaking reform associated with distributions $F(\cdot)$ and $G(\cdot)$. Let $\bar{\rho}$ and $\bar{\rho}'$ be their counterparts for the overall likelihood of reform. For a given $\hat{\theta}$, then, $\rho > \rho'$ and $\bar{\rho} > \bar{\rho}'$ if $F(\cdot)$ first order stochastically dominates $G(\cdot)$.*

Now, we discuss the intuition and implications of these results. Part 1 of Proposition 3 states that the low-type politician conducts more reforms when the public holds a less favourable prior assessment, or when the proportion of capable politicians in the population is smaller. A more favourable prior assessment increases a politician's loss from a failed reform, which consequently weakens his incentive to reform. By contrast, a less favourable prior assessment strengthens his incentive to take risks, because it implies a smaller loss from a failed reform but a larger gain from an accidental success. This is then interpreted as the *pressure to prove oneself* phenomenon.

Part 1 of Proposition 3 further shows that less reform would take place overall if the politician in office is more likely to be a capable one (when the public has a more favourable initial assessment of the politician's talent, or when there is a higher proportion of capable politicians). Note that

$$\frac{\partial \bar{\rho}}{\partial \alpha} = [1 - F(\hat{\theta})][1 - \rho^* + (1 - \alpha)\frac{\partial \rho^*}{\partial \alpha}]. \tag{5}$$

Two competing forces come into play when $\alpha$ is higher. On the one hand, since the low-type politician reforms less than the high-type politician, more reform would be expected if the politician in office is more likely to be a high-type one. This effect is depicted by the term $(1-\rho^*)$. On the other hand, a larger $\alpha$ leads the low-type politician to reform less frequently. This effect is embodied by the term $(1 - \alpha)\frac{\partial \rho^*}{\partial \alpha}$. Our analysis shows that the latter effect always dominates the former.

This result yields an empirically testable hypothesis: when there is a smaller proportion of capable politicians in the population or when the public holds a more pessimistic prior view, more reform can be expected. Conversely, the public observes less reform when politicians have a better initial reputation. This conclusion is drawn without knowledge of the true type of the politician, which is his private information and is unverifiable.

This phenomenon can be witnessed in a wide variety of contexts, and our result sheds light on it. Young or less established individuals are usually seen as being more progressive and opposed to the status quo, in contrast to senior or more established individuals, who usually behave more prudently and conservatively. A famous example is the "'Young' Turks" reform movement, which agitated against the Ottoman Empire in the early 20th century,

building a rich tradition of dissent and paving the foundation of modern Turkey.[18] The term "Young Turks" today represents progressive individuals who are eager to bring about widespread change.

Our analysis may also account for the controversial and seemingly "imprudent" move of Lee Hsien Loong, the current prime minister of Singapore, in 2004. Lee visited Taiwan and demonstrated conspicuously his interests in mediating Sino-Taiwan relation. This move deviated from the long-standing policy paradigm set by Lee Kuan Yew, his father and the founding father of Singapore, who in the 1990s stopped mediating relations between the two parties and committed to Singapore's "non-involvement" in cross-Strait affairs. Lee Hsien Loong's visit caused turbulence to the country's relations with both (mainland) China and Taiwan. Political commentators regarded his visit as the demonstration of lack of "diplomacy and delicacy" in handling international relations. Various possible explanations of his actions have been offered. A rationale, however, can be found in light of the *pressure to prove oneself* phenomenon. Lee's visit coincided with the official confirmation of his prime ministership. An analogy was often drawn between his rise and dynastic succession. This move can be plausibly interpreted as an attempt to establish his credibility and independence.

Part 2 of Proposition 3 states that a low-type politician would mimic his high-type counterpart less often when the latter is more capable. The logic of this result is as follows. When the high-type politician has a more accurate signal, the public is more likely to attribute an unsuccessful reform to a low-type politician. This effect unambiguously increases the low-type politician's costs for carrying out reforms, thereby leading him to reform less often. This result is interpreted as the *tough act to follow* phenomenon.

The distribution of $\theta$ does not qualitatively alter the main prediction of our analysis, but it quantitatively affects the equilibrium behaviour. Part 3 of Proposition 3 describes its effect on $\rho^*$. A stochastically dominant distribution implies that the probability mass is shifted upward. Hence, favourable reform proposals are more likely to be realized. Given the better prospects for reform, the public would then believe that a no-reform outcome is more likely to be caused by the politician's lack of talent, instead of a lack of opportunities (a lower realization of $\theta$). The public's assessment of the politician's ability is therefore lowered when they observe no reform, and this "forces" the low-type politician to reform more often. This result yields interesting welfare implications, which are discussed later in this paper.

[18]The Young Turks originated from the secret societies of progressive and modernist university students and military cadets, who advocated reformation of the Ottoman administration and promoted social and political changes against the monarchy. The Young Turk revolution re-established the constitutional era in Turkey in 1908. As a nationalist party, the Young Turks dominated Turkey's domestic politics thereafter for an entire decade.

# 4 Institution Design

In our equilibrium analysis, the legislature abides by the "constitution" $\hat{\theta}$, which limits the politician's scope of discretion and sets the standard for admissible reform. Based on these results, we turn to the investigation of the optimal institution $\hat{\theta}^*$ that maximizes social welfare.

In our model, a higher $\hat{\theta}$ represents a more conservative rule that grants less authority to the politician; while a lower $\hat{\theta}$ represents a more liberal rule that is more permissive of reform. As aforementioned, the society may expect a gain from the reform that is undertaken by the high-type politician (when $\theta$ is sufficiently high), while it always expects a loss from the reform that is undertaken by the low type. A trade-off is triggered when a more conservative rule is adopted. By restricting reform, it reduces the damage from the latter on the one hand, while it also reduces the gain from the former on the other.

Under an arbitrary threshold rule $\hat{\theta}$, the social welfare in this equilibrium can be written as a function

$$W = \underbrace{\alpha \int_{\hat{\theta}}^{\theta_2} [\theta - 4(1-q)] f(\theta) d\theta}_{W_1} + \underbrace{(1-\alpha)\rho^* \int_{\hat{\theta}}^{\theta_2} (\theta - 2) f(\theta) d\theta}_{W_2}. \tag{6}$$

By implementing a proposal of value $\theta \geq \hat{\theta}$, the high-type politician contributes an expected outcome of $\theta - 4(1-q)$, while the low-type generates a loss of $(\theta - 2)$. The term $W_1$ thus represents the overall net gain from the reform that is undertaken by the high type; while the term $W_2$ depicts the overall (negative) gain from the inefficient reform that is undertaken by the low type. Clearly, the optimal rule $\hat{\theta}^*$ must exceed $4(1-q)$.

Consider an arbitrary reform proposal with a value $\theta \in [\hat{\theta}, \theta_2]$. The *ex ante* expected outcome of this proposal under a threshold rule $\hat{\theta}$ is given by

$$E(y|\, \theta, \hat{\theta}) = \alpha[\theta - 4(1-q)] + (1-\alpha)\rho^*(\theta - 2),$$

which, for a given $\hat{\theta}$, strictly increases with $\theta$. Define $\underline{\rho} \equiv \lim_{\hat{\theta} \uparrow \theta_2} \rho^*$. We have the following.

**Lemma 1.** *Whenever*

$$\frac{(1-\alpha)\underline{\rho}}{\alpha} < \frac{\theta_2 - 4(1-q)}{2 - \theta_2}, \tag{7}$$

*there exists a unique* $\hat{\theta}^0 \in (4(1-q), \theta_2)$ *that solves*

$$E(y|\, \hat{\theta}, \hat{\theta}) = \alpha[\hat{\theta} - 4(1-q)] + (1-\alpha)\rho^*(\hat{\theta} - 2) = 0.$$

*Further,* $\hat{\theta}^0$ *exhibits the following property: for any* $\hat{\theta} \in [-\theta_1, \theta_2]$,

$$E(y|\hat{\theta}, \hat{\theta}) \gtreqless 0 \text{ if and only if } \hat{\theta} \gtreqless \hat{\theta}^0. \tag{8}$$

Suppose that an arbitrary threshold rule $\hat{\theta}$ is being implemented. The expression of $E(y|\,\hat{\theta},\hat{\theta})$ depicts the expected outcome from a "marginal" reform proposal, i.e. the proposal with a value of exactly $\hat{\theta}$. The property of $\hat{\theta}^0$ demonstrated by (8) yields interesting implications. Specifically, the threshold rule $\hat{\theta}^0$ can be viewed as a natural benchmark. If the prevailing rule $\hat{\theta}$ is less conservative than $\hat{\theta}^0$, it must admit "bad" reform: reform with a value in $[\hat{\theta},\hat{\theta}^0)$ would be allowed, which yields negative expected outcome.[19] In contrast, if the prevailing rule $\hat{\theta}$ imposes more restrictions than $\hat{\theta}^0$, it must thwart otherwise "good" reform: reform with a value in $[\hat{\theta}^0,\hat{\theta})$ would be prohibited, which would otherwise yield a positive expected outcome. Hence, a threshold rule $\hat{\theta}^0$, by its very definition, can be labeled as a "neutral" cutoff: it perfectly rules out "bad" reform, while it does not thwart otherwise beneficial reform. Hence, is $\hat{\theta}^0$ the optimal cutoff $\hat{\theta}^*$ that maximizes social welfare? If not, then would the optimal institution be more conservative or less conservative, i.e., does the optimum require $\hat{\theta}^* < \hat{\theta}^0$ or $\hat{\theta}^* > \hat{\theta}^0$? Our analysis yields the following result.

**Proposition 4.** *A unique socially optimal cutoff $\hat{\theta}^* \in (\hat{\theta}^0,\theta_2)$ exists if and only if (7) is satisfied; otherwise, the public prefers no reform at all, i.e., $\hat{\theta}^* = \theta_2$.*

This proposition states that a unique optimal threshold exists, and the optimum $\hat{\theta}^*$ must exceed $\hat{\theta}^0$ whenever $\hat{\theta}^0$ exists. The welfare maximizing institutional rule requires a more stringent standard than $\hat{\theta}^0$. In order to understand its logic, let us now analyze the marginal impact of an increase in $\hat{\theta}$ on social welfare. Taking the first order derivative of (6) with respect to $\hat{\theta}$ yields

$$\frac{dW}{d\hat{\theta}} = f(\hat{\theta})\left\{\begin{array}{l} \underbrace{-\alpha[\hat{\theta}-4(1-q)]}_{a} - \underbrace{(1-\alpha)\rho^*(\hat{\theta}-2)}_{b} \\ \qquad +(1-\alpha)\underbrace{\dfrac{d\rho^*/d\hat{\theta}}{f(\hat{\theta})}\displaystyle\int_{\hat{\theta}}^{\theta_2}(\theta-2)f(\theta)d\theta}_{c} \end{array}\right\}. \tag{9}$$

An increase in $\hat{\theta}$ affects $W$ through three venues. First, it reduces the beneficial reform that is undertaken by the high type, and therefore decreases the gains from this source. This loss is shown by the term $a$, which is negative whenever $\hat{\theta} > 4(1-q)$. Second, a higher cutoff $\hat{\theta}$ (directly) reduces the *expected* loss from the inefficient reform that is undertaken by the low type. This (direct) effect is embodied by the term $b$. Third, it exercises an indirect effect. A higher cutoff further leads the low-type politician to refrain from undertaking reform for any given $\theta \geq \hat{\theta}$ (because $d\rho^*/d\hat{\theta} < 0$ by Proposition 3), which further reduces the loss from his inefficient reform. This positive (indirect) effect is depicted by the term $c$.

[19] Based on the definition of $\hat{\theta}^0$, under the threshold $\hat{\theta} < \hat{\theta}^0$, even a reform with a value that is higher than $\hat{\theta}^0$ may still incur an expected loss.

The decomposition of $dW/d\hat{\theta}$ demonstrates that $\hat{\theta}^0$ is never the optimal threshold. When $\hat{\theta} = \hat{\theta}^0$, the sum of the first two terms ($a$ and $b$) simply boils down to $-E(y|\,\hat{\theta}^0, \hat{\theta}^0)$, and is equal to zero by the definition of $\hat{\theta}^0$. The last term ($c$), however, remains positive. It implies that $W$ can be further increased by raising $\hat{\theta}$ from $\hat{\theta}^0$: although a more conservative threshold would deter otherwise beneficial reform, it further deters the detrimental reform that is undertaken by the low type by decreasing $\rho^*$. Our analysis reveals that a unique optimal threshold $\hat{\theta}^* > \hat{\theta}^0$ exists. By implementing this conservative rule, the reduced loss more than compensates for the sacrificed gain from those otherwise efficient reforms with value in $(\hat{\theta}^0, \hat{\theta}^*)$. In conclusion, the social optimum must require a sufficiently cautious attitude towards potential reform, despite it inhibiting seemingly beneficial reform.

Reform can be permitted, i.e., $\hat{\theta}^* < \theta_2$, if and only if condition (7) is met. Because $\underline{\rho}$ decreases with $\alpha$ (by Proposition 1), the left hand side of (7) strictly decreases with $\alpha$. Hence, this condition is more likely to be met with a larger $\alpha$, i.e., the presence of a higher proportion of high-talent politicians in the population. When the talent required for successful reform is very scarce, the public would not expect sufficient gain from reform. The public then prefer no reform at all. Similarly, the condition is more likely to be met with a larger $q$. In other words, reform is socially beneficial only when the success of reform is sufficiently likely.

These arguments further lead to more general conclusions on the impact of $\alpha$ and $q$ on the properties of $\hat{\theta}^* \in (\hat{\theta}^0, \theta_2)$. A greater $\alpha$ or $q$ always allows for less restriction on the politician's activities, which is formally stated as follows.

**Proposition 5.** *The socially optimal cutoff $\hat{\theta}^*$ decreases with $\alpha$ and $q$.*

## "Optimism" Requires More Caution

Proposition 3 demonstrates that the equilibrium behaviour depends on the properties of the distribution of $\theta$. We now discuss its impact on the optimal threshold rule $\hat{\theta}^*$. To allow for a handy and informative analysis, we restrict our attention to an example where the value of reform follows a uniform distribution

$$F(\theta) = \frac{\theta + \theta_1}{\theta_2 + \theta_1}$$

and the high-talent politician receives a perfect signal with $q = 1$, which allows for a closed form solution to $\rho^*$.

An increase in $\theta_2$ implies that the probability mass of the distribution is shifted upward, high-valued reform proposals are more likely to occur, and more beneficial opportunities can be expected. The environment thus seems to favor more reform. Before we examine its impact on the socially optimal institutional rule, let us examine its welfare implications in an arbitrary equilibrium with a fixed $\hat{\theta}$. Figure 1 testifies to a non-monotonic relationship

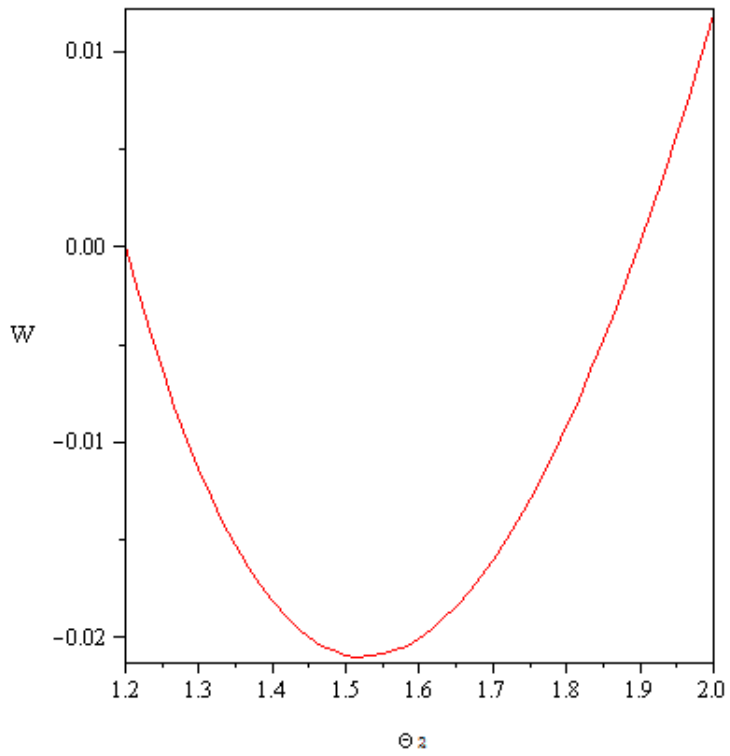


Figure 1: An example that demonstrates the non-monotonic effect of $\theta_2$ on social welfare ($\theta_1 = 1.1$, $q = 1.0$, $\alpha = 0.2$, and $\hat{\theta} = 1.2$).

between social welfare and $\theta_2$ when $\hat{\theta}$ is given. The society may not be better off when more opportunities are available. The logic can be seen in Proposition 3: for a given cutoff $\hat{\theta}$, a stochastically dominant distribution of $\theta$ forces the low type to reform more, which increases the loss from his inefficient reform.

The ambiguous welfare implication compels us to further look into its implications on the socially optimal institution $\hat{\theta}^*$. The implications of a higher $\theta_2$ on the social optimum $\hat{\theta}^*$ can be ambiguous as well. On the one hand, low-valued reform proposals would emerge less often, and cause less damage, which encourages a more liberal rule to reap more benefits from reform. On the other hand, it could demand a more conservative rule in order to further discipline the low type. Our analysis leads to the following proposition.

**Proposition 6.** *The socially optimal cutoff point for reform, $\hat{\theta}^*$, strictly increases with $\theta_2$.*

We find that when the probability mass of the uniform distribution is shifted upward, i.e., when more opportunities for reform can be expected, it unambiguously lifts the optimal cutoff $\hat{\theta}^*$. That is, a more favourable environment requires additional caution and a more conservative institutional rule.

# 5 Discussion and Extensions

In this section, we further discuss the main features of our model and its possible extensions. We first examine the relations and differences between this study and studies by Majumdar and Mukand (2004), Suurmond, Swank, and Visser (2004), and Chen (2010). As discussed in Section 1, our paper differs from these papers in both focuses and settings.

In the models of Majumdar and Mukand (2004) and Suurmond, Swank, and Visser (2004) (among others), the performance of the risky project is depicted by a binary indicator (success or failure) alone. The likelihood of its success is pre-determined and a more capable agent can discover the pre-determined "suitability" of the project more precisely. In contrast, we assume that (1) the performance of reform is a continuous measure, depending on the quality of both the project *per se* and the implementation by the politician; (2) the politician's ability determines the quality of *ex-post* implementation; and (3) the quality of the project is random.

Our model is closer to Chen (2010), as in both studies, the likelihood of success of the risky project is assumed to depend on the agent's ability. However, our setup also differs from Chen's (2010) in several aspects. First, in Chen's model (2010), the likelihood of success depends on a random variable, whose realization is observable only to the agent, while the principal's payoffs from success or failure of the risky project are prefixed. In contrast, in our model, the public's payoffs from a successful or failed reform are dependent on a random variable, and the probabilities of success or failure are affected by the politician's ability alone. Second, in Chen's (2010) model, even if the risky project fails, the agent's reputation is still higher than that from choosing the safe project.[20] In our model, a politician's reputation declines after a failed reform, compared to maintaining the status quo.

These diverse modelling approaches serve the differing research focuses of these studies, such that these papers complement each other. Our setup enriches the analysis in two aspects. First, it enables the analysis of institution design. An extensive trade-off is involved in determining the proper level of institutional conservatism. Second, a comparative static analysis may be performed on the probability distribution of the value of reform, which sheds further light on the equilibrium behaviour and the corresponding design of welfare-maximizing institution.

Our analysis has been limited to a stylized setting for the sake of expositional efficiency and mathematical tractability. It, however, leaves open many possibilities for extensions and variations. A simple and straightforward alternative is to consider "transparency" as an institutional element in our context. Suppose that the public is able to learn the "counterfactual", i.e., the true realization of $\theta$, when the status quo is maintained. The analysis on

[20] In her model, the agent is prevented from choosing the risky project only by monetary incentives.

our basic setting (Proof of Proposition 1) immediately implies that both types would reform with probability one whenever $\theta \geq \hat{\theta}$. The low-type politician strictly prefers undertaking the reform, as his type would be otherwise completely revealed.[21] A "wrong kind of transparency" can loom large in our context as well: a non-transparent environment allows the low type to better hide his type, and leads him to refrain from inefficient risk-taking. Our paper echoes the conclusion of Andrea Prat (2005) in this aspect, although transparency and its negative effects appear in differing contexts. In Section 5.3, we explore the ramifications of transparency of an alternative form in a different context.

In the remainder of this paper, we discuss briefly a few possible extensions to our basic framework. Although these extensions would not yield predictions that fundamentally depart from our main results, they may spawn richer comparative statics that further improve our understanding of this issue.

## 5.1 Repeated Policy Choices

The analysis presented above may also be extended to a dynamic setting, where the politician makes his policy choice repeatedly. Let us consider a two-stage game, where the politician is allowed to decide whether to reform in both periods. Although space restrictions prevent us from presenting the detailed analysis from the extended setting, our basic analysis yields immediate implications. The "pressure to prove oneself" result points toward the following prediction: a politician who has failed in the past is more likely to take radical action in the future. Past failure lowers his public ratings which then make it more lucrative for him to pursue accidental success in the future. As a result, it can be demonstrated that in the first period of the two-stage game, the low-type politician reforms less often than he would in a static setting. To put it intuitively, his failure not only jeopardizes his reputation for the current period, but also "forces" him to risk more in future, thereby further reducing his payoff. Such concerns compel him to refrain from taking risks in the earlier stage.

## 5.2 When the Politician Values Policy Performance

Our model can be extended to allow the payoff of the politician to depend on the realized outcome of his policy choice. Assume that the politician cares not only about his reputation payoff, but also receives utility from the output of his policy choice. His objective function is written generically as

$$u(y, i) = \delta \Pr(t = H | \, y, i) + (1 - \delta) y, \tag{10}$$

with $\delta \in [0, 1]$.

[21] In this case, he cannot pool himself with a high type who has no opportunity ($\theta < \hat{\theta}$).

A smaller $\delta$ implies that the politician is subject to weaker reputation concerns. The model boils down to the first best benchmark when $\delta$ reduces to zero, while it approximates our original model when $\delta$ approaches one. We characterize the equilibrium of the game with a given $\hat{\theta}$ in the following proposition. For expositional efficiency, we consider only the case of $\hat{\theta} \geq 4(1-q)$.

**Remark 1.** *In equilibrium, the low-type politician reforms with a positive probability if and only if $\delta$ is sufficiently large, i.e., when the condition $\frac{\delta}{1-\delta}\frac{1-\alpha}{\alpha F(\hat{\theta})+(1-\alpha)} > 2-\theta_2$ is met. Under this condition, the high-type politician reforms with probability one for all $\theta \geq \hat{\theta}$, and there exists a unique cutoff $\overline{\theta}_L \in [\hat{\theta}, \theta_2)$, such that the low type reforms with a probability $\rho_L^*(\theta|\,\hat{\theta}) \in (0,1)$ for all $\theta > \overline{\theta}_L$, with $\rho_L^*(\theta|\,\hat{\theta})$ strictly increasing with $\theta$.*

The proof is similar to that for Proposition 1.[22] The equilibrium of the game ultimately depends on the size of $\delta$. When the politician's utility also depends on the actual output $y$, the politician bears additional loss from his unsuccessful reform, which may discourage a low-type politician from undertaking reform. It comes as no surprise that the low-type politician must reform less often than he would in our basic model. When $\delta$ is sufficiently small, the low-type politician may even completely abstain from undertaking the reform. However, whenever nontrivial reputation concerns are present, i.e. $\frac{\delta}{1-\delta}\frac{1-\alpha}{\alpha F(\hat{\theta})+(1-\alpha)} > 2-\theta_2$, the equilibrium behaviour resembles that in the basic model: the low type mimics his high-type counterpart, in spite of the additional loss from his more likely failure. However, he plays a strictly monotone equilibrium strategy with $\rho_L^*(\theta)$ strictly increasing with $\theta$: his failure would cost less, if he implements a more valuable proposal. We further present the following remark.

**Remark 2.** *Under nontrivial reputation concerns, i.e. $\frac{\delta}{1-\delta}\frac{1-\alpha}{\alpha F(\hat{\theta})+(1-\alpha)} > 2-\theta_2$, the low-type politician reforms less often when a higher $\hat{\theta}$ is in place. That is, the cutoff $\overline{\theta}_L$ strictly increases with $\hat{\theta}$, and $\rho_L^*(\theta|\,\hat{\theta})$ strictly decreases with $\hat{\theta}$ for all $\theta \in [\overline{\theta}_L, \theta_2]$.*

Similar to Proposition 2, Remark 2 demonstrates that the low-type politician would reform less when a more stringent standard is in place. The positive effect of a higher $\hat{\theta}$, which was discussed in Section 4, remains in the extended setting (where the politician also cares about the actual outcome $y$): with nontrivial reputation concerns, a more restrictive institution deters inefficient reforms conducted by low-type politicians and reduces the damage from them.

[22] We omit it for brevity but it is available from the author upon request.

## 5.3 Imperfect Observation of Reform Outcome

So far, we have assumed that the outcome of reform is revealed to the public post-reform. To make our model more realistic, we may allow for the possibility that the public does not perfectly observe the outcome of the reform. Let us consider the situation where there is a post-reform evaluation that, with probability $\eta$, allows the public to find out how well a particular reform fared. Otherwise, the evaluation discovers nothing. The parameter $\eta$ can be interpreted as the transparency of the political environment – how well the public can monitor the policy performance of a politician. Now, the equilibrium condition for the low-type politician to mix between reform and the status quo can be rewritten as

$$\mu^0 = \eta\left(\frac{1}{2}\mu^s + \frac{1}{2}\mu^f\right) + (1-\eta)\mu^n,$$

where $\mu^n$ is the reputation of the politician if the post-reform evaluation does not discover the performance of the reform. Rewriting it in the manner of (3) gives

$$\frac{1}{1+\lambda(\alpha)A} = \eta\left(\frac{1}{2}\cdot\frac{1}{1+\lambda(\alpha)B} + \frac{1}{2}\cdot\frac{1}{1+\lambda(\alpha)C}\right) + (1-\eta)\frac{1}{1+\lambda(\alpha)\rho}, \tag{11}$$

where as before

$$\lambda(\alpha) = \frac{1-\alpha}{\alpha},\ A = 1+(1-\rho)\kappa(\hat{\theta}),\ \kappa(\hat{\theta}) = \frac{1-F(\hat{\theta})}{F(\hat{\theta})},\ B = \frac{\frac{1}{2}\rho}{q},\ C = \frac{\frac{1}{2}\rho}{1-q}.$$

Note that a low-type politician's expected reputation is lower if the performance of the reform is discovered than if it is not. The reason is that when the reform outcome is discovered, it is more difficult for the low-type politician to disguise himself as a high-type politician, relative to the situation when the public does not discover the actual outcome, because he is more likely to fail in reform than the high type.

In fact, by undertaking the reform with any positive probability $\rho$, the reputation of the low-type politician if no discovery is made, $\mu^n$, is strictly higher than his reputation from choosing the status quo, $\mu^0$. Therefore, if $\eta$ is sufficiently small, e.g. $\eta \to 0$, the low-type politician reforms with probability one as long as the high-type politician does, which results in a "pooling equilibrium." To put it intuitively, the low-type politician would be punished less severely if his failure is less likely to be found out. This incentivizes him to risk more. However, if $\eta$ is sufficiently large, then again we have the semi-separating equilibrium as in our previous analysis, in which the low-type politician mimics his high-type counterpart with a probability of $\rho^* \in (0,1)$. "Transparency" has the effect of deterring a low-type politician from taking inefficient reform. To summarize, we conclude with the following.

**Remark 3.** *The low-type politician's probability of reform, $\rho^*$, is non-increasing in $\eta$, $\alpha$, and $q$.*

This extension does not cause qualitative changes to our equilibrium predictions. The above remark shows that all the results of Proposition 3 continue to hold. That is, the "pressure to prove oneself" and the "tough act to follow" phenomena continue to exist.

It would be interesting to investigate how the optimal institution varies with parameters in this setup. Though we strongly conjecture that our results in the section on institution design will continue to hold, the complexity of the calculations prevents us from drawing a definite conclusion. However, we will pursue these results in future research.

# 6 Concluding Remarks

In this paper, we study a politician's incentive to implement reform when his true ability is privately known but he is concerned about the public's perception of his abilities. The politician thus chooses his policy to maximize his reputation payoff. We find that a high-talent politician always attempts to reform as much as possible, which compels his low-talent counterpart to mimic with a positive probability. Socially inefficient reform therefore results. Further, we explore the socially optimal level of empowerment in the presence of such reputation concerns of the politician. We find that the social optimum can be achieved only if the prevailing institutional rule embodies proper conservatism and deters some otherwise efficient reform.

# 7 Appendix: Proofs

## 7.1 Proof of Proposition 1

### 7.1.1 Divinity Criterion

We first formally translate the notion of the Divinity Criterion into our context. Suppose that in a hypothetical equilibrium, in which there exists $\theta \in [\hat{\theta}, \theta_2]$, with $\rho_t(\theta) = 0$, $\forall t \in \{L, H\}$. Suppose that an unexpected reform with a value $\theta \geq \hat{\theta}$ takes place. The public infers from its outcome the value of $\theta$. The public forms a set of beliefs $\phi_\theta \equiv \{\widetilde{\rho}_H(\theta), \widetilde{\rho}_L(\theta)\}$, where $\widetilde{\rho}_t(\theta)$ specifies the probability of a type-$t$ politician to undertake this reform. Given this conjecture, a type-$t$ politician, when deviating, has a payoff

$$\mu_t(\theta; \phi_\theta) = q_t \times \frac{\alpha\widetilde{\rho}_H(\theta)q}{\alpha\widetilde{\rho}_H(\theta)q + \frac{1}{2}(1-\alpha)\widetilde{\rho}_L(\theta)} + (1-q_t) \times \frac{\alpha\widetilde{\rho}_H(\theta)(1-q)}{\alpha\widetilde{\rho}_H(\theta)(1-q) + \frac{1}{2}(1-\alpha)\widetilde{\rho}_L(\theta)}.$$

Let $\mu_t^*$ denote the payoff of a type-$t$ politician in the equilibrium. Further define $\Phi_\theta^t \equiv \{\phi_\theta |\, \mu_t(\theta; \phi_\theta) > \mu_t^*\}$. We then have the following.

**Definition 1.** Under Divinity Criterion, the out-of-equilibrium belief $\phi_\theta$ satisfies:

$$\widetilde{\rho}_t(\theta) \geq \widetilde{\rho}_{t'}(\theta) \text{ if } \Phi_\theta^{t'} \subset \Phi_\theta^t, \text{ with } t \in \{H, L\} \text{ and } t \neq t'.$$

We claim the following results.

**Claim 1.** *Suppose that in a hypothetical equilibrium, in which there exists $\theta \in [\hat{\theta}, \theta_2]$, with $\rho_t(\theta) = 0$, $\forall t \in \{L, H\}$. Then $\Phi_\theta^L \subset \Phi_\theta^H$.*

*Proof.* Consider a hypothetical deviation of a reform with value $\theta$. Define $\widetilde{\alpha} \equiv \frac{\alpha\widetilde{\rho}_H(\theta)}{\alpha\widetilde{\rho}_H(\theta)+(1-\alpha)\widetilde{\rho}_L(\theta)}$. The high type, if deviates, has an *ex ante* expected payoff

$$\begin{aligned}\mu_H(\theta;\widetilde{\alpha}) &= q \times \frac{\widetilde{\alpha}q}{\widetilde{\alpha}q + \frac{1}{2}(1-\widetilde{\alpha})} + (1-q) \times \frac{\widetilde{\alpha}(1-q)}{\widetilde{\alpha}(1-q) + \frac{1}{2}(1-\widetilde{\alpha})} \\ &= q \times \frac{1}{1 + \frac{1}{2\widetilde{\alpha}q}(1-\widetilde{\alpha})} + (1-q) \times \frac{1}{1 + \frac{1}{2\widetilde{\alpha}(1-q)}(1-\widetilde{\alpha})}.\end{aligned}$$

She has an incentive to deviate if and only if $\pi_H(\theta) - \mu^0 \geq 0$. The low type, by contrast, has an *ex ante* expected payoff

$$\mu_L(\theta;\widetilde{\alpha}) = \frac{1}{2} \times \frac{1}{1 + \frac{1}{2\widetilde{\alpha}q}(1-\widetilde{\alpha})} + \frac{1}{2} \times \frac{1}{1 + \frac{1}{2\widetilde{\alpha}(1-q)}(1-\widetilde{\alpha})}.$$

She has an incentive to deviate if and only if $\pi_L(\theta) - \mu^0 \geq 0$. Because $\frac{1}{1+\frac{1}{2\widetilde{\alpha}q}(1-\widetilde{\alpha})} > \frac{1}{1+\frac{1}{2\widetilde{\alpha}(1-q)}(1-\widetilde{\alpha})}$, we see that $\mu_H(\theta) - \mu^0 > 0$ whenever $\mu_L(\theta) - \mu^0 \geq 0$. It implies that the high type is always more likely to deviate by undertaking an expected reform than the low type. ∎

Claim 1 demonstrates that the high type always benefits more from reform.

### 7.1.2 Equilibrium Characterization

**Claim 2.** *There exists no equilibrium with $\theta \in [\hat{\theta}, \theta_2]$, and $\rho_t(\theta) = 0$, $\forall t \in \{L, H\}$.*

*Proof.* The out-of-equilibrium belief must require $\widetilde{\alpha} \geq \alpha$ to reflect the result of Claim 1. We now prove $\mu_H(\theta) > \alpha$. To see this, observe that

$$\begin{aligned}\mu_H(\theta;\widetilde{\alpha}) &= q \times \frac{\widetilde{\alpha}q}{\widetilde{\alpha}q + \frac{1}{2}(1-\widetilde{\alpha})} + (1-q) \times \frac{\widetilde{\alpha}(1-q)}{\widetilde{\alpha}(1-q) + \frac{1}{2}(1-\widetilde{\alpha})} \\ &> [\widetilde{\alpha}q + \frac{1}{2}(1-\widetilde{\alpha})] \times \frac{\widetilde{\alpha}q}{\widetilde{\alpha}q + \frac{1}{2}(1-\widetilde{\alpha})} \\ &\quad + [\widetilde{\alpha}(1-q) + \frac{1}{2}(1-\widetilde{\alpha})] \times \frac{\widetilde{\alpha}(1-q)}{\widetilde{\alpha}(1-q) + \frac{1}{2}(1-\widetilde{\alpha})} \\ &= \widetilde{\alpha} > \alpha,\end{aligned}$$

where we have used the fact that $q > 1/2$. Given such a belief, the high type must deviate when $\theta$ is realized, because his expected payoff $\mu_H(\theta) > \alpha > \mu^0$. The original equilibrium cannot be sustained by a belief system that satisfies Divinity. ∎

**Claim 3.** *For any $\theta \in [\hat{\theta}, \theta_2]$, in an equilibrium,*

*1. if the low type reforms with positive probability, the high type must reform with probability one;*

*2. if the high type does not reform, the low type would not reform with positive probability;*

*3. $\rho_H(\theta) > 0$, $\forall \theta \in [\hat{\theta}, \theta_2]$.*

*Proof.* When the politician does not undertake a reform, his expected payoff $\mu^0$ is independent of his type. Whenever he reform, the expected payoff for a high type is $q\mu^s + (1-q)\mu^f$, which is higher than that for the low type, $\frac{1}{2}\mu^s + \frac{1}{2}\mu^f$. Hence, if we have $\frac{1}{2}\mu^s + \frac{1}{2}\mu^f \geq \mu^0$, then $q\mu^s + (1-q)\mu^f > \mu^0$. Further, if we have the high type choose not to reform, i.e. $q\mu^s + (1-q)\mu^f \leq \mu^0$, then $\frac{1}{2}\mu^s + \frac{1}{2}\mu^f < \mu^0$.

Suppose that there exists $\theta \in [\hat{\theta}, \theta_2]$ with $\rho_H(\theta) > 0$. By Claim 2 $\rho_L(\theta) > 0$. Contradiction. ∎

**Claim 4.** *In equilibrium, the politician must play monotone strategy, such that $\rho_t(\theta)$ must be nondecreasing with $\theta$ for $\theta \in [\hat{\theta}, \theta_2]$, $\forall t \in \{L, H\}$.*

*Proof.* Suppose that there exist $\theta, \theta' \in [\hat{\theta}, \theta_2]$, with $\theta > \theta'$ and $\rho_t(\theta) < \rho_t(\theta')$. Recall the definition of $q_t$. We must have

$$q_t\mu^s(\theta) + (1-q_t)\mu^f(\theta) \leq q_t\mu^s(\theta') + (1-q_t)\mu^f(\theta'),$$

which is written as

$$\begin{aligned} & q_t\frac{\alpha q\rho_H(\theta)}{\alpha q\rho_H(\theta) + (1-\alpha)\frac{1}{2}\rho_L(\theta)} + (1-q_t)\frac{\alpha(1-q)\rho_H(\theta)}{\alpha(1-q)\rho_H(\theta) + (1-\alpha)\frac{1}{2}\rho_L(\theta)} \\ \leq \; & q_t\frac{\alpha q\rho_H(\theta')}{\alpha q\rho_H(\theta') + (1-\alpha)\frac{1}{2}\rho_L(\theta')} + (1-q_t)\frac{\alpha(1-q)\rho_H(\theta')}{\alpha(1-q)\rho_H(\theta') + (1-\alpha)\frac{1}{2}\rho_L(\theta')}. \end{aligned} \tag{12}$$

By Claim 2, $\rho_H(\cdot)$ cannot be zero. The condition is further rewritten as

$$\begin{aligned} & q_t[\frac{\alpha q}{\alpha q + (1-\alpha)\frac{1}{2}\frac{\rho_L(\theta)}{\rho_H(\theta)}} - \frac{\alpha q}{\alpha q + (1-\alpha)\frac{1}{2}\frac{\rho_L(\theta')}{\rho_H(\theta')}}] \\ \leq \; & (1-q_t)[\frac{\alpha(1-q)}{\alpha(1-q) + (1-\alpha)\frac{1}{2}\frac{\rho_L(\theta')}{\rho_H(\theta')}} - \frac{\alpha(1-q)}{\alpha(1-q) + (1-\alpha)\frac{1}{2}\frac{\rho_L(\theta)}{\rho_H(\theta)}}], \end{aligned}$$

which requires $\frac{\rho_L(\theta)}{\rho_H(\theta)} \geq \frac{\rho_L(\theta')}{\rho_H(\theta')}$.

Suppose $\rho_L(\theta), \rho_L(\theta') > 0$, then by Claim 3 $\rho_H(\theta) = \rho_H(\theta') = 1$. Then we have $\rho_L(\theta) \geq \rho_L(\theta')$. Contradiction.

Suppose $\rho_L(\theta) = \rho_L(\theta') = 0$. In that case, undertaking reform gives a payoff of one, which is not an equilibrium, as the low type must deviate. Contradiction. ∎

**Claim 5.** *$\rho_H(\theta) = 1$ and $\rho_L(\theta) > 0$, $\forall \theta \in [\hat{\theta}, \theta_2]$.*

*Proof.* We claim that whenever the high type chooses reform with a positive probability, the low type must do so as well. We have shown that whenever both types choose reform with positive probability, the high type's probability of reform is one and therefore at least as high as the low type's. Therefore, the overall probability for the low type to choose the status quo, $P_{0L}$, is weakly higher than that for the high type, $P_{0H}$. Thus, if the low type chooses the status quo, his reputation is $\mu^0 = \frac{\alpha P_{0H}}{\alpha P_{0H} + (1-\alpha) P_{0L}} \leq \alpha$.

However, if he deviates and undertakes reform, he is believed to be a high type with probability one if $q < 1$. If $q = 1$, his payoff depends on the public's off-equilibrium belief when reform fails. However, he succeeds with probability $\frac{1}{2}$, and the resulting expected payoff still exceeds $\alpha$. Therefore, it cannot be that the low type always chooses the status quo when the high type chooses reform.

By Claim 3, $\rho_H(\theta) > 0$, $\forall \theta \in [\hat{\theta}, \theta_2]$. Then $\rho_L(\theta) > 0$, $\forall \theta \in [\hat{\theta}, \theta_2]$. Again, by Claim 3, $\rho_H(\theta) = 1$, $\forall \theta \in [\hat{\theta}, \theta_2]$. ∎

### 7.1.3 Proof of Proposition 1

*Proof.* We now determine the low-type politician's probability of reform for a proposal with value $\theta$, which we denote by $\rho(\theta)$ to economize on notation. By (3), if the politician maintains the status quo, his payoff is

$$\begin{aligned} \mu^0 &= \frac{\alpha F(\hat{\theta})}{\alpha F(\hat{\theta}) + (1-\alpha) F(\hat{\theta}) + (1-\alpha) \int_{\hat{\theta}}^{\theta_2} [1 - \rho(\theta)] f(\theta) d\hat{\theta}} \\ &= \frac{\alpha}{\alpha + (1-\alpha) \frac{F(\hat{\theta}) + \int_{\hat{\theta}}^{\theta_2} [1-\rho(\theta)] f(\theta) d\theta}{F(\hat{\theta})}}. \end{aligned} \tag{13}$$

Note that it does not depend on $\theta$. On the other hand, if the low-type politician undertakes the reform, his payoff is given by

$$\begin{aligned} \mu_L(\theta) &= \frac{1}{2} \cdot \frac{q\alpha f(\theta)}{q\alpha f(\theta) + \frac{1}{2}(1-\alpha)\rho(\theta) f(\theta)} + \frac{1}{2} \cdot \frac{(1-q)\alpha f(\theta)}{(1-q)\alpha f(\theta) + \frac{1}{2}(1-\alpha)\rho(\theta) f(\theta)} \\ &= \frac{1}{2} \cdot \frac{\alpha}{\alpha + \frac{\frac{1}{2}(1-\alpha)\rho(\theta)}{q}} + \frac{1}{2} \cdot \frac{\alpha}{\alpha + \frac{\frac{1}{2}(1-\alpha)\rho(\theta)}{1-q}}. \end{aligned} \tag{14}$$

If the low-type plays a completely mixed strategy, $\rho(\theta) \in (0, 1)$, we need to equate (13) and (14), which implies that $\rho(\theta)$ must be a constant $\rho$ regardless of the value $\theta$. Consequently, in equilibrium,

$$\frac{\alpha}{\alpha + (1-\alpha) \frac{F(\hat{\theta}) + (1-\rho^*)[1 - F(\hat{\theta})]}{F(\hat{\theta})}} = \frac{1}{2} \cdot \frac{\alpha}{\alpha + (1-\alpha) \frac{\frac{1}{2}\rho^*}{q}} + \frac{1}{2} \cdot \frac{\alpha}{\alpha + (1-\alpha) \frac{\frac{1}{2}\rho^*}{1-q}}, \tag{15}$$

which we may rewrite as

$$\frac{1}{1+\lambda(\alpha)A} = \frac{1}{2}\cdot\frac{1}{1+\lambda(\alpha)B} + \frac{1}{2}\cdot\frac{1}{1+\lambda(\alpha)C}, \tag{16}$$

where

$$\lambda(\alpha) = \frac{1-\alpha}{\alpha},\ A = 1+(1-\rho)\kappa(\hat{\theta}),\ \kappa(\hat{\theta}) = \frac{1-F(\hat{\theta})}{F(\hat{\theta})},\ B = \frac{\frac{1}{2}\rho}{q},\ C = \frac{\frac{1}{2}\rho}{1-q}.$$

This is the same equation as (3). The expression $\lambda(\alpha)$ is the likelihood ratio of the low type versus the high type, $\kappa(\hat{\theta})$ is the likelihood ratio of reform having good prospects versus bad prospects, and $A$, $B$, and $C$ are respectively the likelihood ratios of the low type not reforming, having a successful reform, and having a failed reform versus the high type obtaining each outcome. Consider the equilibrium condition (16). Note that its $LHS$ is $\mu^0$ and its $RHS$ is $\mu_L$. When $\rho = 0$, $\mu^0 \leq \alpha$, while $\mu_L = 1$ as $B = C = 0$. Therefore, $\mu^0 < \mu_L$. By contrast, when $\rho = 1$, $\mu^0 = \alpha$ as $A = 1$, and $\mu_{1L} < \alpha$, which can be seen from the fact that when $\rho = 1$, $\alpha\mu_H + (1-\alpha)\mu_L = \alpha$ must hold while $\mu_L < \mu_H$. Therefore, $\mu^0 > \mu_L$.

Both the $RHS$ and $LHS$ of (16) are continuous in $\rho$. Furthermore, it is straightforward to show that the $LHS$ strictly increases with $\rho$, while the $RHS$ strictly decreases with $\rho$. Hence, we conclude that there must exist a unique $\rho^* \in (0,1)$ that solves (16). ∎

## 7.2 Proof of Proposition 2

*Proof.* We first establish the following.

**Claim 6.** *In equilibrium, $\rho^*$ strictly decreases with $\hat{\theta}$.*

Rewrite the equilibrium condition as

$$g(\rho^*, \alpha, q, \hat{\theta}) \equiv [1+(1-\rho^*)\kappa(\hat{\theta})] - \frac{\rho^*[\lambda(\alpha)\rho^*+1]}{4q(1-q)+\lambda(\alpha)\rho} = 0 \tag{17}$$

with $\kappa(\hat{\theta}) \equiv \frac{1-F(\hat{\theta})}{F(\hat{\theta})}$. When $\hat{\theta}$ increases, $\kappa(\hat{\theta})$ must decrease, which causes $g(\rho^*, \alpha, q, \hat{\theta})$ to decrease. Further, as we have shown in the proof for previous results, $g(\rho^*, \alpha, q, \hat{\theta})$ strictly decreases with $\rho^*$. By the implicit function theorem, we establish that when $\hat{\theta}$ increases, $\rho^*$ must decrease.

We then use the result to establish the main claim of Proposition 3. Recall that the equilibrium is defined by the equation

$$\underbrace{\frac{\alpha}{1+\frac{(1-\alpha)(1-\rho^*)[1-F(\hat{\theta})]}{F(\hat{\theta})}}}_{\mu^0} = \frac{1}{2}[\underbrace{\frac{\alpha}{\alpha+(1-\alpha)\frac{\frac{1}{2}\rho^*}{q}}}_{\mu^s} + \underbrace{\frac{\alpha}{\alpha+(1-\alpha)\frac{\frac{1}{2}\rho^*}{1-q}}}_{\mu^f}].$$

The politician in office receives a payoff $\mu^0$ when he maintains the status quo. He receives a payoff $\mu^s$ when he successfully implements a reform and $\mu^f$ when he fails. In any equilibrium with a given $\hat{\theta}$, the type-$t$ politician receives a payoff

$$\mu_t = \begin{cases} q_t\mu^s + (1-q_t)\mu^f, & \text{for } \theta \geq \hat{\theta}; \\ \mu^0, & \text{for } \theta < \hat{\theta} \end{cases}.$$

Hence, in this equilibrium, the expected payoff of a type-$t$ politician is given by

$$E(\mu_t) = \mu^0 F(\hat{\theta}) + [q_t\mu^s + (1-q_t)\mu^f][1-F(\hat{\theta})].$$

First, we claim that when $\hat{\theta}$ increases, $E(\mu_H)$ and $E(\mu_L)$ change in opposite directions. Therefore, the first part of the proposition implies the second part. This claim is an implication of the fact $\alpha E(\mu_H) + (1-\alpha)E(\mu_L) = \alpha$, or $E(\mu_H) = 1 - \lambda(\alpha)E(\mu_L)$.

Now, we prove the first part of the proposition. For a low-type politician, $E(\mu_L) = \mu^0$ because $\mu^0 = \frac{1}{2}\mu^s + \frac{1}{2}\mu^f$. Hence, we need only verify $\frac{d\mu^0}{d\hat{\theta}} > 0$. Define

$$H(\rho^*, \hat{\theta}) = \underbrace{\frac{\alpha}{1 + \frac{(1-\alpha)(1-\rho^*)[1-F(\hat{\theta})]}{F(\hat{\theta})}}}_{\mu^0} - \frac{1}{2}[\underbrace{\frac{\alpha}{\alpha + (1-\alpha)\frac{\frac{1}{2}\rho^*}{q}}}_{\mu^s} + \underbrace{\frac{\alpha}{\alpha + (1-\alpha)\frac{\frac{1}{2}\rho^*}{1-q}}}_{\mu^f}].$$

We have

$$\frac{d\mu^0}{d\hat{\theta}} = \frac{\partial\mu^0}{\partial\hat{\theta}} + \frac{\partial\mu^0}{\partial\rho^*}\cdot\frac{\partial\rho^*}{\partial\hat{\theta}} = \frac{\partial\mu^0}{\partial\hat{\theta}} + \frac{\partial\mu^0}{\partial\rho^*}\cdot[-\frac{\partial H(\rho^*,\hat{\theta})}{\partial\hat{\theta}} \Big/ \frac{\partial H(\rho^*,\hat{\theta})}{\partial\rho^*}].$$

Because $\frac{\partial H(\rho^*,\hat{\theta})}{\partial\hat{\theta}} = \frac{\partial\mu^0}{\partial\hat{\theta}}$, we then have $\frac{d\mu^0}{d\hat{\theta}} = \frac{\partial\mu^0}{\partial\hat{\theta}}[1 - \frac{\partial\mu^0}{\partial\rho^*} \Big/ \frac{\partial H(\hat{\theta},\rho^*)}{\partial\rho^*}]$. We must have $1 - \frac{\partial\mu^0}{\partial\rho^*} \Big/ \frac{\partial H(\hat{\theta},\rho^*)}{\partial\rho^*} > 0$ because $\frac{\partial H(\rho^*,\hat{\theta})}{\partial\rho^*} = \frac{\partial\mu^0}{\partial\rho^*} - \frac{1}{2}(\frac{\partial\mu^s}{\partial\rho^*} + \frac{\partial\mu^f}{\partial\rho^*})$, while $\frac{\partial\mu^0}{\partial\rho^*} > 0$, $\frac{\partial\mu^s}{\partial\rho^*}, \frac{\partial\mu^f}{\partial\rho^*} < 0$. ∎

## 7.3 Proof of Proposition 3

*Proof.* **Part 1** Consider the equilibrium condition (16). We have shown above that the left hand side of (16) is increasing in $\rho^*$ and the right hand side decreasing in $\rho^*$. Note that $A$, $B$, and $C$ do not contain $\alpha$ in their expressions. Thus, we may write

$$\frac{\partial(LHS - RHS) \text{ of (16)}}{\partial\alpha} = -\frac{1}{\alpha^2}\left[-\frac{A}{(1+\lambda(\alpha)A)^2} + \frac{1}{2}\cdot\frac{B}{(1+\lambda(\alpha)B)^2} + \frac{1}{2}\cdot\frac{C}{(1+\lambda(\alpha)C)^2}\right].$$

We want to evaluate the above derivative at the value of $\rho$ that satisfies (16). Observe that $0 < B < C$ as $q \geq 3/4 > 1/2$, we may conclude then $B < A < C$ based on (16). From (16), we obtain

$$\frac{A}{1+\lambda(\alpha)A} = \frac{1}{2}\cdot\frac{B}{1+\lambda(\alpha)B} + \frac{1}{2}\cdot\frac{C}{1+\lambda(\alpha)C}.$$

Therefore,

$$\begin{aligned}
&\frac{1}{2}\cdot\frac{B}{\left(1+\lambda(\alpha)B\right)^2}+\frac{1}{2}\cdot\frac{C}{\left(1+\lambda(\alpha)C\right)^2}\\
=\;&\frac{A}{1+\lambda(\alpha)A}\left[\frac{\frac{B}{1+\lambda(\alpha)B}}{\frac{B}{1+\lambda(\alpha)B}+\frac{C}{1+\lambda(\alpha)C}}\cdot\frac{1}{1+\lambda(\alpha)B}+\frac{\frac{C}{1+\lambda(\alpha)C}}{\frac{B}{1+\lambda(\alpha)B}+\frac{C}{1+\lambda(\alpha)C}}\cdot\frac{1}{1+\lambda(\alpha)C}\right].
\end{aligned}$$

The expression in the brackets is a convex combination of $\frac{1}{1+\lambda(\alpha)B}$ and $\frac{1}{1+\lambda(\alpha)C}$. Since $0<B<C$, the former is larger, but the coefficient on the former is smaller than $\frac{1}{2}$. Using (16), we have

$$\frac{1}{2}\cdot\frac{B}{\left(1+\lambda(\alpha)B\right)^2}+\frac{1}{2}\cdot\frac{C}{\left(1+\lambda(\alpha)C\right)^2}<\frac{A}{\left(1+\lambda(\alpha)A\right)^2}.$$

Hence, at the value of $\rho$ that satisfies (16),

$$\frac{\partial(LHS-RHS)\text{ of (16)}}{\partial\alpha}>0.$$

Thus, by the implicit function theorem, the probability of reform by the low type, $\rho^*$, is decreasing in $\alpha$, the probability of high type.

Next, we verify the comparative statics of $\bar{\rho}$. Because $\frac{\partial\rho^*}{\partial\alpha}<0$, we only need to show $\left|(1-\alpha)\frac{\partial\rho^*}{\partial\alpha}\right|+\rho^*>1$. We have

$$\begin{aligned}
&\left|(1-\alpha)\frac{\partial\rho^*}{\partial\alpha}\right|+\rho^*\\
=\;&\frac{(1-\alpha)}{\alpha^2}\frac{\frac{\rho^{*2}[1-[4q(1-q)]}{[4q(1-q)+\lambda(\alpha)\rho^*]^2}}{[\kappa(\hat{\theta})+1+\frac{4q(1-q)[1-[4q(1-q)]}{[4q(1-q)+\lambda(\alpha)\rho^*]^2}]}+\rho^*
\end{aligned}$$

Rearranging the equilibrium condition leads to

$$\begin{aligned}
(1-\rho^*)\kappa(\hat{\theta}) &= \frac{\rho^*(\lambda(\alpha)\rho^*+1)}{4q(1-q)+\lambda(\alpha)\rho^*}-1\\
&= \frac{\rho^*(\lambda(\alpha)\rho^*+1)-4q(1-q)-\lambda(\alpha)\rho^*}{4q(1-q)+\lambda(\alpha)\rho^*}\\
&= \frac{\lambda(\alpha)\rho^{*2}+\rho^*-\lambda(\alpha)\rho^*-4q(1-q)}{4q(1-q)+\lambda(\alpha)\rho^*},
\end{aligned}$$

which yields $\kappa(\hat{\theta})=\frac{\lambda(\alpha)\rho^{*2}+\rho^*-\lambda(\alpha)\rho^*-4q(1-q)}{[4q(1-q)+\lambda(\alpha)\rho^*](1-\rho^*)}$, and therefore

$$\begin{aligned}
\kappa(\hat{\theta})+1 &= \frac{\lambda(\alpha)\rho^{*2}+\rho^*-\lambda(\alpha)\rho^*-4q(1-q)+[4q(1-q)+\lambda(\alpha)\rho^*](1-\rho^*)}{[4q(1-q)+\lambda(\alpha)\rho^*](1-\rho^*)}\\
&= \frac{\rho^*[1-4q(1-q)]}{[4q(1-q)+\lambda(\alpha)\rho^*](1-\rho^*)}.
\end{aligned}$$

Hence,

$$\begin{aligned} & [\kappa(\hat{\theta}) + 1 + \frac{4q(1-q)[1-4q(1-q)]}{[4q(1-q)+\lambda(\alpha)\rho^*]^2}] \\ = & \frac{\rho[1-4q(1-q)]}{[4q(1-q)+\lambda(\alpha)\rho^*](1-\rho^*)} + \frac{4q(1-q)[1-4q(1-q)]}{[4q(1-q)+\lambda(\alpha)\rho^*]^2} \\ = & \frac{1-4q(1-q)}{[4q(1-q)+\lambda(\alpha)\rho^*]^2(1-\rho^*)}[4q(1-q)+\lambda(\alpha)\rho^{*2}]. \end{aligned}$$

We then obtain

$$\begin{aligned} & \left|(1-\alpha)\frac{\partial \rho^*}{\partial \alpha}\right| + \rho^* \\ = & \frac{(1-\alpha)}{\alpha^2} \cdot \frac{\frac{\rho^{*2}[1-4q(1-q)]}{[4q(1-q)+\lambda(\alpha)\rho^*]^2}}{\frac{1-4q(1-q)}{[4q(1-q)+\lambda(\alpha)\rho^*]^2(1-\rho^*)}[4q(1-q)+\lambda(\alpha)\rho^{*2}]} + \rho^* \\ = & \frac{(1-\alpha)}{\alpha^2} \cdot \frac{(1-\rho^*)\rho^{*2}}{[4q(1-q)+\lambda(\alpha)\rho^{*2}]} + \rho^*. \end{aligned}$$

For our purpose, we only need to show $\frac{(1-\alpha)}{\alpha^2} \cdot \frac{\rho^{*2}}{[4q(1-q)+\lambda(\alpha)\rho^{*2}]} > 1$. Rewrite it as $\frac{(1-\alpha)}{\alpha^2} \cdot \frac{\rho^{*2}}{[4q(1-q)+\lambda(\alpha)\rho^{*2}]} = \frac{1}{\alpha} \cdot \frac{\lambda(\alpha)\rho^{*2}}{[4q(1-q)+\lambda(\alpha)\rho^{*2}]} = \frac{1}{\alpha} \cdot \frac{1}{[\frac{4q(1-q)}{\lambda(\alpha)\rho^{*2}}+1]}$. Hence, it suffices to show $\frac{1}{[\frac{4q(1-q)}{\lambda(\alpha)\rho^{*2}}+1]} > \alpha$. We claim $\frac{1}{[\frac{4q(1-q)}{\lambda(\alpha)\rho^{*2}}+1]} > \frac{1}{2} > \alpha$, i.e., $4q(1-q) < \lambda(\alpha)\rho^{*2}$. To show that, recall the equilibrium condition $1+(1-\rho^*)m] = \frac{\rho^*(\lambda(\alpha)\rho^*+1)}{4q(1-q)+\lambda(\alpha)\rho^*}$, which implies $\frac{\rho^*(\lambda(\alpha)\rho^*+1)}{4q(1-q)+\lambda(\alpha)\rho^*} > 1 \Leftrightarrow \rho^*(\lambda(\alpha)\rho^*+1) > 4q(1-q)+\lambda(\alpha)\rho^* \Leftrightarrow \lambda(\alpha)\rho^{*2}+\rho^* > 4q(1-q)+\lambda(\alpha)\rho^*$. Because $\lambda(\alpha) > 1$, $\lambda(\alpha)\rho^{*2} > 4q(1-q)$ must hold.

**Part 2** Recall the equilibrium condition (17). Since $q \geq \frac{3}{4}$, $G(\rho^*, \alpha, q, \hat{\theta})$ is decreasing with $q$. Further,

$$\frac{\partial g(\rho^*, \alpha, q, \hat{\theta})}{\partial \rho^*} = -\left[\kappa(\hat{\theta}) + 1 + \frac{4q(1-q)[1-[4q(1-q)]^2]}{[4q(1-q)+\lambda\rho^*]^2}\right] < 0.$$

Recall that $\kappa(\hat{\theta}) = [1-F(\hat{\theta})]/F(\hat{\theta})$. We then obtain $\frac{d\rho^*}{dq} = -\frac{\frac{\partial g(\rho^*,q)}{\partial q}}{\frac{\partial g(\rho^*,q)}{\partial \rho^*}} < 0$. That $\bar{\rho}$ is decreasing in $q$ is an immediate consequence by its definition in (4).

**Part 3** Consider the equilibrium condition (16). Since $F$ first order stochastically dominates $G$, we have $F(\hat{\theta}) < G(\hat{\theta})$. This implies that for any given $\rho$, LHS of (16) for $F$ is lower than that for $G$, since $\kappa(\hat{\theta})$ is larger for $F$ than for $G$. As we have shown above, LHS of (16) strictly increases with $\rho$, while RHS strictly decreases. Thus, only if $\rho > \rho'$ can make (16) hold for both distributions. From this, the definition of $\bar{\rho}$ in (4), and the assumption that $F$ first order stochastically dominates $G$, we can immediately see that $\bar{\rho} > \bar{\rho}'$. ∎

## 7.4 Proof of Lemma 1

*Proof.* Consider the value of

$$E(y|\,\hat{\theta},\hat{\theta}) = \alpha[\hat{\theta} - 4(1-q)] + (1-\alpha)\rho^*(\hat{\theta} - 2).$$

When $\hat{\theta} = 4(1-q)$, it must be negative. When $\hat{\theta}$ approaches $\theta_2$, we have its value approach $\alpha[\theta_2 - 4(1-q)] + (1-\alpha)\underline{\rho}(\theta_2 - 2)$, which is positive if and only if $\frac{(1-\alpha)\underline{\rho}}{\alpha} < \frac{\theta_2 - 4(1-q)}{2-\theta_2}$. Further recall that $E(y|\,\theta,\hat{\theta})$ strictly increases with both $\theta$ and $\hat{\theta}$. There must exist a unique $\hat{\theta}^0$ that solves the equation. ∎

## 7.5 Lemma 2 and Its Proof

Because $f(\hat{\theta}) > 0$ for all $\hat{\theta} \in [-\theta_1, \theta_2]$, the sign of (9) is the same as that of $\frac{dW}{d\hat{\theta}} \diagup f(\hat{\theta})$. For our purpose, it suffices to explore $\frac{dW}{d\hat{\theta}} \diagup f(\hat{\theta})$. We then establish the following lemma.

**Lemma 2.** *The expression $\frac{dW}{d\hat{\theta}} \diagup f(\hat{\theta})$ strictly decreases with $\hat{\theta}$.*

*Proof.* Recall that the equilibrium condition (17) with a given $\hat{\theta}$ can be written as

$$g(\rho^*, \kappa(\hat{\theta})) = [1 + (1-\rho^*)\kappa(\hat{\theta})] - \frac{\rho^*(\lambda\rho^* + 1)}{4q(1-q) + \lambda\rho^*} = 0,$$

where $\kappa(\hat{\theta}) = [1 - F(\hat{\theta})]/F(\hat{\theta})$. Hence, we have $\frac{\partial g(\rho^*, \kappa(\hat{\theta}))}{\partial \kappa(\hat{\theta})} = \lambda(1-\rho^*)$. Because $\frac{\partial g(\rho^*, \kappa(\hat{\theta}))}{\partial \rho^*} = -[\kappa(\hat{\theta}) + 1 + \frac{4q(1-q)[1-[4q(1-q)]^2]}{[4q(1-q)+\lambda\rho]^2}] < 0$, we must have

$$\frac{d\rho^*}{d\kappa(\hat{\theta})} = \frac{1-\rho^*}{\kappa(\hat{\theta}) + 1 + \frac{4q(1-q)[1-[4q(1-q)]^2]}{[4q(1-q)+\lambda\rho^*]^2}},$$

and therefore $\frac{d\rho^*}{d\hat{\theta}} \diagup f(\hat{\theta}) = -\frac{1-\rho^*}{[\kappa(\hat{\theta})+1+\frac{4q(1-q)[1-[4q(1-q)]^2]}{[4q(1-q)+\lambda\rho^*]^2}][F(\hat{\theta})]^2}$.

We now claim $-\frac{d\rho^*}{d\hat{\theta}} \diagup f(\hat{\theta})$ strictly decreases with $\hat{\theta}$. We have

$$\frac{d[-\frac{d\rho^*}{d\hat{\theta}} \diagup f(\hat{\theta})]}{d\hat{\theta}} = \frac{\left[\begin{array}{c} -\frac{d\rho^*}{d\hat{\theta}}[\kappa(\hat{\theta}) + 1 + \frac{4q(1-q)[1-[4q(1-q)]^2]}{[4q(1-q)+\lambda\rho^*]^2}][F(\hat{\theta})]^2 \\ -(1-\rho^*)\frac{d\{[\kappa(\hat{\theta})+1+\frac{4q(1-q)[1-[4q(1-q)]^2]}{[4q(1-q)+\lambda\rho^*]^2}][F(\hat{\theta})]^2\}}{d\hat{\theta}} \end{array}\right]}{\{[\kappa(\hat{\theta}) + 1 + \frac{4q(1-q)[1-[4q(1-q)]^2]}{[4q(1-q)+\lambda\rho^*]^2}][F(\hat{\theta})]^2\}^2}.$$

Note that $-\frac{d\rho^*}{d\hat{\theta}}[\kappa(\hat{\theta}) + 1 + \frac{4q(1-q)[1-[4q(1-q)]^2]}{[4q(1-q)+\lambda\rho^*]^2}][F(\hat{\theta})]^2 = (1-\rho^*)f(\hat{\theta})$. We then only need to prove $\frac{d\{[\kappa(\hat{\theta})+1+\frac{4q(1-q)[1-[4q(1-q)]^2]}{[4q(1-q)+\lambda\rho^*]^2}][F(\hat{\theta})]^2\}}{d\hat{\theta}} > f(\hat{\theta})$. Rewrite $[\kappa(\hat{\theta})+1+\frac{4q(1-q)[1-[4q(1-q)]^2]}{[4q(1-q)+\lambda\rho^*]^2}][F(\hat{\theta})]^2$ as $F(\hat{\theta}) + \frac{4q(1-q)[1-[4q(1-q)]^2]}{[4q(1-q)+\lambda\rho^*]^2}[F(\hat{\theta})]^2$. When $\hat{\theta}$ increases, both $\frac{4q(1-q)[1-[4q(1-q)]^2]}{[4q(1-q)+\lambda\rho^*]^2}$ and $F(\hat{\theta})$ strictly increases. Hence, $\frac{d\{\frac{4q(1-q)[1-[4q(1-q)]^2]}{[4q(1-q)+\lambda\rho^*]^2}][F(\hat{\theta})]^2\}}{d\hat{\theta}} > 0$. Furthermore, $\frac{dF(\hat{\theta})}{d\hat{\theta}} = f(\hat{\theta})$. We then establish our claim. ∎

## 7.6 Proof of Proposition 4

*Proof.* If $\frac{(1-\alpha)\underline{\rho}}{\alpha} \geq \frac{\theta_2-4(1-q)}{2-\theta_2}$, then $\hat{\theta}^0$ does not exist. Any reform with a value $\theta < \theta_2$ must lead to negative expected outcome. Hence, no reform is *ex ante* beneficial, which implies $\hat{\theta}^* = \theta_2$. If $\frac{(1-\alpha)\underline{\rho}}{\alpha} < \frac{\theta_2-4(1-q)}{2-\hat{\theta}}$, then $\hat{\theta}^0$ exists. $\frac{dW}{d\hat{\theta}} \diagup f(\hat{\theta})\Big|_{\hat{\theta}=\hat{\theta}^0} > 0$, but $\frac{dW}{d\hat{\theta}} \diagup f(\hat{\theta})\Big|_{\hat{\theta}=\theta_2} < 0$ (because $\frac{(1-\alpha)\underline{\rho}}{\alpha} < \frac{\theta_2-4(1-q)}{2-\hat{\theta}}$), then there must exist a unique $\hat{\theta}^* \in (\hat{\theta}^0, \theta_2)$ that solves $\frac{dW}{d\hat{\theta}} \diagup f(\hat{\theta}) = 0$. ∎

## 7.7 Proof of Proposition 5

*Proof.* Suppose that an interior optimum with $\hat{\theta}^* \in (0, \theta_2)$ exists. Define $k \equiv [-\frac{d\rho^*}{d\hat{\theta}} \diagup f(\hat{\theta})]$. Then the optimal condition is

$$\upsilon \equiv \alpha[\hat{\theta} - 4(1-q)] + (1-\alpha)\rho^*(\hat{\theta}-2) - (1-\alpha)k\int_{\hat{\theta}}^{\theta_2}(2-\theta)f(\theta)d\theta = 0. \tag{18}$$

Apparently, $\frac{d\upsilon}{d\hat{\theta}} = -\frac{d\frac{\frac{dW}{d\hat{\theta}}}{f(\hat{\theta})}}{d\hat{\theta}} > 0$. We now claim $\frac{d\upsilon}{d\alpha} > 0$. Taking first order derivative of $\upsilon$ yields

$$\begin{aligned}\frac{d\upsilon}{d\alpha} &= [\hat{\theta} - 4(1-q)] - \rho^*(\hat{\theta}-2) + (1-\alpha)\frac{d\rho^*}{d\alpha}(\hat{\theta}-2) \\ &\quad + k\int_{\hat{\theta}}^{\theta_2}(2-\theta)f(\theta)d\theta - (1-\alpha)\frac{\partial k}{\partial \alpha}\int_{\hat{\theta}}^{\theta_2}(2-\theta)f(\theta)d\theta.\end{aligned}$$

It suffices to show $k$ strictly decreases with $\alpha$ and $q$. Recall by the proofs of previous results:

$$-\frac{d\rho^*}{d\alpha} = \frac{\frac{\rho^{*2}[1-[4q(1-q)]}{[4q(1-q)+\lambda\rho^*]^2}}{[\kappa(\hat{\theta}) + 1 + \frac{4q(1-q)[1-[4q(1-q)]^2]}{[4q(1-q)+\lambda\rho^*]^2}]} \cdot \left|\frac{d\lambda(\alpha)}{d\alpha}\right|.$$

Note $-\frac{d\frac{d\rho^*}{d\hat{\theta}}}{d\alpha} = -\frac{d\frac{d\rho^*}{d\alpha}}{d\hat{\theta}}$. Hence, we now evaluate $-\frac{d\rho^*}{d\alpha}$ with respect to $\hat{\theta}$. We first rearrange it as

$$\begin{aligned}-\frac{d\rho^*}{d\alpha} &= \frac{\frac{\rho^{*2}[1-[4q(1-q)]}{[4q(1-q)+\lambda\rho^*]^2}}{[\kappa(\hat{\theta}) + 1 + \frac{4q(1-q)[1-[4q(1-q)]^2]}{[4q(1-q)+\lambda\rho^*]^2}]} \cdot \left|\frac{d\lambda(\alpha)}{d\alpha}\right| \\ &= \frac{(1-\rho^*)}{[\kappa(\hat{\theta}) + 1 + \frac{4q(1-q)[1-[4q(1-q)]^2]}{[4q(1-q)+\lambda\rho^*]^2}]} \cdot [1-[4q(1-q)] \\ &\quad \cdot \frac{\rho^{*2}}{1-\rho^*} \cdot \frac{1}{[4q(1-q)+\lambda\rho^*]^2}.\end{aligned}$$

By the proof of Part 1 of Proposition 2, $\frac{(1-\rho^*)}{[\kappa(\hat{\theta})+1+\frac{4q(1-q)[1-[4q(1-q)]^2]}{[4q(1-q)+\lambda\rho^*]^2}]}$ decreases with $\hat{\theta}$. We claim $\frac{\rho^{*2}}{1-\rho^*} \cdot \frac{1}{[4q(1-q)+\lambda\rho^*]^2}$ also decreases with $\hat{\theta}$. Evaluate it with respect to $\hat{\theta}$ yields

$$\frac{\rho^*(2-\rho^*)\frac{d\rho^*}{d\hat{\theta}}}{(1-\rho^*)^2} \cdot \frac{1}{[4q(1-q)+\lambda\rho^*]^2} + \frac{\rho^{*2}}{1-\rho^*} \cdot \frac{-2\lambda\frac{d\rho^*}{d\hat{\theta}}}{[4q(1-q)+\lambda\rho^*]^3}.$$

Because $\frac{d\rho^*}{d\hat{\theta}} < 0$, we need to show $(2-\rho^*)[4q(1-q)+\lambda\rho^*]-2\lambda\rho^*(1-\rho^*) > 0$, which is obvious because $(2-\rho^*)[4q(1-q)+\lambda\rho^*]-2\lambda\rho^*(1-\rho^*) = (2-\rho^*)[4q(1-q)+\lambda\rho^*]-\lambda\rho^*(2-2\rho^*)$, and $2-\rho^* > 2-2\rho^*$.

We further claim $\hat{\theta}^*$ decreases with $q$. To show that, we have to prove $\frac{d\upsilon}{dq} > 0$. We have

$$\frac{d\upsilon}{dq} = 4\alpha + (1-\alpha)\frac{d\rho^*}{dq}(\hat{\theta}-2) - (1-\alpha)\frac{dk}{dq}\int_{\hat{\theta}}^{\theta_2}(2-\theta)f(\theta)d\theta.$$

It would suffice to show $\frac{dk}{dq} < 0$. We use the same technique as above. We have

$$-\frac{d\rho^*}{dq} = \frac{\frac{4(2q-1)\rho^*(\lambda\rho^*+1)}{[4q(1-q)+\lambda\rho^*]^2}}{[\kappa(\bar{\theta})+1+\frac{4q(1-q)[1-[4q(1-q)]^2]}{[4q(1-q)+\lambda\rho^*]^2}]}.$$

We then claim $-\frac{\partial^2\rho^*}{\partial q\partial\hat{\theta}} < 0$. Rewrite $-\frac{d\rho^*}{dq}$ as

$$-\frac{d\rho^*}{dq} = \frac{1-\rho^*}{[\kappa(\hat{\theta})+1+\frac{4q(1-q)[1-[4q(1-q)]^2]}{[4q(1-q)+\lambda\rho^*]^2}]}\cdot\frac{1}{1-\rho^*}\cdot\frac{4(2q-1)\rho^*(\lambda\rho^*+1)}{[4q(1-q)+\lambda\rho^*]^2}.$$

Because $\frac{1-\rho^*}{[\kappa(\hat{\theta})+1+\frac{4q(1-q)[1-[4q(1-q)]^2]}{[4q(1-q)+\lambda\rho^*]^2}]}$ and $\frac{1}{1-\rho^*}$ decreases with $\hat{\theta}$, we only need to show $\frac{\rho^*(\lambda\rho^*+1)}{[4q(1-q)+\lambda\rho^*]^2}$ decreases with $\hat{\theta}$. Taking first order derivative of it with respect to $\hat{\theta}$ yields

$$\begin{aligned}
\frac{d\frac{\rho^*(\lambda\rho^*+1)}{[4q(1-q)+\lambda\rho^*]^2}}{d\hat{\theta}} &= \frac{\begin{bmatrix}(2\lambda\rho^*+1)\frac{d\rho^*}{d\hat{\theta}}[4q(1-q)+\lambda\rho^*]^2(1-\rho^*)\\ -2\rho^*(\lambda\rho^*+1)(1-\rho^*)[4q(1-q)+\lambda\rho^*]\lambda\frac{d\rho^*}{d\hat{\theta}}\\ +\rho^*(\lambda\rho^*+1)[4q(1-q)+\lambda\rho^*]^2\frac{d\rho^*}{d\hat{\theta}}\end{bmatrix}}{(1-\rho^*)^2[4q(1-q)+\lambda\rho^*]^4}\\
&= \frac{\frac{d\rho^*}{d\hat{\theta}}}{[4q(1-q)+\lambda\rho^*]^3}\times\begin{bmatrix}(2\lambda\rho^*+1)[4q(1-q)+\lambda\rho^*](1-\rho^*)\\ -2\lambda\rho^*(\lambda\rho^*+1)(1-\rho^*)\\ +\rho^*(\lambda\rho^*+1)[4q(1-q)+\lambda\rho^*]\end{bmatrix}.
\end{aligned}$$

By Proposition 2, $\frac{d\rho^*}{d\hat{\theta}} < 0$. Hence, it remains to verify that the item in bracket is positive. This is obvious because

$$\begin{aligned}
&\begin{bmatrix}(2\lambda\rho^*+1)[4q(1-q)+\lambda\rho^*](1-\rho^*)\\ -2\lambda\rho^*(\lambda\rho^*+1)(1-\rho^*)\\ +\rho^*(\lambda\rho^*+1)[4q(1-q)+\lambda\rho^*]\end{bmatrix}\\
> \;& \lambda\rho^*\left[(2\lambda\rho^*+1)(1-\rho^*)-2(\lambda\rho^*+1)(1-\rho^*)+\rho^*(\lambda\rho^*+1)\right]\\
= \;& \lambda\rho^*[-\rho^*+\rho^*(\lambda\rho^*+1)] > 0. \;\blacksquare
\end{aligned}$$

## 7.8 Proof of Proposition 6

*Proof.* We examine how a higher upper support $\theta_2$ could affect $dW/d\hat{\theta}$ for any given $\hat{\theta}$. When the high-type politician is perfectly informed, a closed form for $\rho^*$ is obtained as

$$\rho^* = 1 - \frac{\alpha}{1-\alpha}F(\hat{\theta}).$$

The first-order derivative of the welfare function is derived as follows

$$\begin{aligned}\frac{dW}{d\hat{\theta}} &= \frac{1}{\theta_2+\theta_1}\left\{\begin{array}{l}-\alpha\hat{\theta}-(1-\alpha)\rho^*(\hat{\theta}-2)\\ +(1-\alpha)\frac{d\rho^*/d\hat{\theta}}{f(\hat{\theta})}\int_{\hat{\theta}}^{\theta_2}(\theta-2)f(\theta)d\theta\end{array}\right\}\\ &= \frac{1}{\theta_2+\theta_1}\left\{\begin{array}{l}-\alpha\hat{\theta}+(1-\alpha)[1-\frac{\alpha(\hat{\theta}+\theta_1)}{(1-\alpha)(\theta_2+\theta_1)}](2-\hat{\theta})\\ +\frac{\alpha(\theta_2-\hat{\theta})}{2}[4-(\theta_2+\hat{\theta})]\end{array}\right\}.\end{aligned}$$

The optimal cutoff $\hat{\theta}^*$ is determined by the equation

$$\upsilon = -\alpha\hat{\theta}+(1-\alpha)[1-\frac{\alpha(\hat{\theta}+\theta_1)}{(1-\alpha)(\theta_2+\theta_1)}](2-\hat{\theta})+\frac{\alpha(\theta_2-\hat{\theta})}{2}[4-(\theta_2+\hat{\theta})]=0.$$

By Lemma 2, $\upsilon$ strictly decreases with $\hat{\theta}$. We only need to show $\frac{\partial\upsilon}{\partial\theta_2}>0$. Apparently, $(1-\alpha)[1-\frac{\alpha(\hat{\theta}+\theta_1)}{(1-\alpha)(\theta_2+\theta_1)}](2-\hat{\theta})$ increases with $\theta_2$. We claim $\frac{\alpha(\theta_2-\hat{\theta})}{2}[4-(\theta_2+\hat{\theta})]$ increases with it as well. Taking first order derivative of $(\theta_2-\hat{\theta})[4-(\theta_2+\hat{\theta})]$ yields $4-(\theta_2+\hat{\theta})-(\theta_2-\hat{\theta})=4-2\theta_2>0$, which completes the proof. ∎

## 7.9 Proof of Remark 2

*Proof.* If the politician maintains the status quo, the output remains zero. His expected payoff is given by

$$u^0=\frac{\alpha F(\hat{\theta})+\alpha\int_{\hat{\theta}}^{\theta_2}[1-\rho_H(\theta)]f(\theta)d\theta}{\left[F(\hat{\theta})+\alpha\int_{\hat{\theta}}^{\theta_2}[1-\rho_H(\theta)]f(\theta)d\theta+(1-\alpha)\int_{\hat{\theta}}^{\theta_2}[1-\rho_L(\theta)]f(\theta)d\theta\right]}.$$

The payoff $u^0$ does not differ from that in the basic model. By contrast, if the politician implements a reform of value $\theta$, he ends up with a payoff

$$u^s(\theta)=\delta\frac{\alpha q\rho_H(\theta)f(\theta)}{\alpha q\rho_H(\theta)f(\theta)+(1-\alpha)\frac{1}{2}\rho_L(\theta)f(\theta)}+(1-\delta)\theta$$

if the reform succeeds; and

$$u^f(\theta)=\delta\frac{\alpha(1-q)\rho_H(\theta)f(\theta)}{\alpha(1-q)\rho_H(\theta)f(\theta)+(1-\alpha)\frac{1}{2}\rho_L(\theta)f(\theta)}+(1-\delta)(\theta-4),$$

if it fails.

Define $\overline{\theta}_L=\min(\theta|\,\rho_L(\theta|\,\hat{\theta})>0)$. Whenever $\rho_L(\theta|\,\hat{\theta})>0$, the equilibrium is determined by the system of equations:

$$\begin{aligned}&\delta\left[\begin{array}{c}\frac{1}{2}\frac{\alpha q}{\alpha q+(1-\alpha)\frac{1}{2}\rho_L(\theta|\hat{\theta})}\\ +\frac{1}{2}\frac{\alpha(1-q)}{\alpha(1-q)+(1-\alpha)\frac{1}{2}\rho_L(\theta|\hat{\theta})}\end{array}\right]+(1-\delta)(\theta-2)\\ =\;&\frac{\alpha F(\hat{\theta})}{\alpha F(\hat{\theta})+(1-\alpha)-(1-\alpha)\int_{\overline{\theta}_L}^{\theta_2}\rho_L(\theta_0|\,\hat{\theta})f(\theta)d\theta},\forall\theta\in[\overline{\theta}_L,\theta_2],\end{aligned}$$

$$\begin{aligned}
&\delta\left[\begin{array}{c}\frac{1}{2}\frac{\alpha q}{\alpha q+(1-\alpha)\frac{1}{2}\rho_L(\theta|\hat{\theta})}\\ +\frac{1}{2}\frac{\alpha(1-q)}{\alpha(1-q)+(1-\alpha)\frac{1}{2}\rho_L(\theta|\hat{\theta})}\end{array}\right]+(1-\delta)(\theta-2)\\
=\ &\delta\left[\begin{array}{c}\frac{1}{2}\frac{\alpha q}{\alpha q+(1-\alpha)\frac{1}{2}\rho_L(\theta'|\hat{\theta})}\\ +\frac{1}{2}\frac{\alpha(1-q)}{\alpha(1-q)+(1-\alpha)\frac{1}{2}\rho_L(\theta'|\hat{\theta})}\end{array}\right]+(1-\delta)(\theta'-2), \forall \theta,\theta'\in[\overline{\theta}_L,\theta_2].
\end{aligned}$$

We first prove that $\rho_L^*(\theta|\,\hat{\theta})$ strictly decreases with $\hat{\theta}$ by contradiction. By the equilibrium condition, for arbitrary $\theta,\theta'\in[\overline{\theta}_L,\theta_2]$, if $\rho_L(\theta|\,\hat{\theta})$ in(de)creases, then $\rho_L(\theta'|\,\hat{\theta})$ must in(de)crease as well.

Suppose that $\hat{\theta}$ drops but $\rho_L(\theta|\,\hat{\theta})$ decreases. To have the low type reform less, the payoff for no reform, i.e. $\frac{\alpha F(\hat{\theta})}{\left[\alpha F(\hat{\theta})+(1-\alpha)-(1-\alpha)\int_{\overline{\theta}_L}^{\theta_2}\rho_L(\theta|\hat{\theta})f(\theta)d\theta\right]}$, must strictly increase. Because $F(\hat{\theta})$ decreases, $(1-\alpha)-(1-\alpha)\int_{\overline{\theta}_L}^{\theta_2}\rho_L(\theta|\,\hat{\theta})f(\theta)d\theta$ must decrease, which requires $\int_{\overline{\theta}_L}^{\theta_2}\rho_L(\theta|\,\hat{\theta})f(\theta)d\theta$ to increase. Let $\hat{\theta}$ drop to $\hat{\theta}$. We consider the following possibilities.

**Case 1: $\overline{\theta}_L$ also increases.** Then $\int_{\overline{\theta}_L}^{\theta_2}\rho_L(\theta|\,\hat{\theta})f(\theta)d\theta$ must decrease. Contradiction.

**Case 2: $\overline{\theta}_L$ decreases.** Let $\overline{\theta}_L$ drop to $\overline{\theta}'_L$. There are altogether four possibilities.

**Case 2.1 $\overline{\theta}_L>\hat{\theta}$ and $\overline{\theta}'_L>\hat{\theta}'$.** Consider $\theta\in(\overline{\theta}'_L,\overline{\theta}_L)$. Before the drop, the low type does not want to undertake the reform for such $\theta$. Hence, we have $\delta+(1-\delta)(\theta-2)<\frac{\alpha F(\hat{\theta})}{\left[\alpha F(\hat{\theta})+(1-\alpha)-(1-\alpha)\int_{\overline{\theta}_L}^{\theta_2}\rho_L(\theta|\hat{\theta})f(\theta)d\theta\right]}$. After the drop, the low type reforms with a positive probability in equilibrium. He receives $\delta\left[\frac{1}{2}\frac{\alpha q}{\alpha q+(1-\alpha)\frac{1}{2}\rho_L(\theta'|\hat{\theta}')}+\frac{1}{2}\frac{\alpha(1-q)}{\alpha(1-q)+(1-\alpha)\frac{1}{2}\rho_L(\theta'|\hat{\theta}')}\right]+(1-\delta)(\theta-2)$, which is less than $\delta+(1-\delta)(\theta-2)$. However, we also have

$$\begin{aligned}
&\delta\left[\frac{1}{2}\frac{\alpha q}{\alpha q+(1-\alpha)\frac{1}{2}\rho_L(\theta'|\,\hat{\theta}')}+\frac{1}{2}\frac{\alpha(1-q)}{\alpha(1-q)+(1-\alpha)\frac{1}{2}\rho_L(\theta'|\,\hat{\theta}')}\right]\\
=\ &\frac{\alpha F(\hat{\theta}')}{\left[\alpha F(\hat{\theta}')+(1-\alpha)-(1-\alpha)\int_{\overline{\theta}'_L}^{\theta_2}\rho_L(\theta|\,\hat{\theta}')f(\theta)d\theta\right]},
\end{aligned}$$

which is required to be greater than $\frac{\alpha F(\hat{\theta})}{\left[\alpha F(\hat{\theta})+(1-\alpha)-(1-\alpha)\int_{\overline{\theta}_L}^{\theta_2}\rho_L(\theta|\hat{\theta})f(\theta)d\theta\right]}$. Contradiction.

**Case 2.2 $\overline{\theta}_L=\hat{\theta}$ and $\overline{\theta}'_L=\hat{\theta}'$.** The payoff from no reform is written as

$$\alpha\diagup\left[\alpha+(1-\alpha)\frac{1-\int_{\overline{\theta}_L}^{\theta_2}\rho_L(\theta_0|\,\hat{\theta})f(\theta)d\theta}{F(\hat{\theta})}\right].$$

Further, when $\hat{\theta}$ drops, the denominator becomes

$$\alpha + (1-\alpha)\frac{1 - \int_{\overline{\theta}_L}^{\theta_2} \rho_L(\theta|\,\hat{\theta}')f(\theta)d\theta - \int_{\hat{\theta}'}^{\hat{\theta}} \rho_L(\theta|\,\hat{\theta}')f(\theta)d\theta}{F(\hat{\theta})}$$

$$= \quad \alpha + (1-\alpha)[\frac{1 - \int_{\overline{\theta}_L}^{\theta_2} \rho_L(\theta|\,\hat{\theta}')f(\theta)d\theta}{F(\hat{\theta}')} - \frac{\int_{\hat{\theta}'}^{\hat{\theta}} \rho_L(\theta|\,\hat{\theta}')f(\theta)d\theta}{F(\hat{\theta}')}].$$

Because $\rho_L(\cdot) < 1$, $\int_{\hat{\theta}'}^{\hat{\theta}} \rho_L(\theta|\,\hat{\theta}')f(\theta)d\theta < F(\hat{\theta}) - F(\hat{\theta}')$, which is defined as $\Delta$. Then $\frac{1-\int_{\overline{\theta}_L}^{\theta_2} \rho_L(\theta|\hat{\theta}')f(\theta)d\theta}{F(\hat{\theta}')} - \frac{\int_{\hat{\theta}'}^{\hat{\theta}} \rho_L(\theta|\hat{\theta}')f(\theta)d\theta}{F(\hat{\theta}')}$ is further rewritten as $\frac{1-\int_{\overline{\theta}_L}^{\theta_2} \rho_L(\theta|\hat{\theta}')f(\theta)d\theta}{F(\hat{\theta})-\Delta} - \frac{\int_{\hat{\theta}'}^{\hat{\theta}} \rho_L(\theta|\hat{\theta}')f(\theta)d\theta}{F(\hat{\theta})-\Delta}$. We now claim $\frac{1-\int_{\hat{\theta}}^{\theta_2} \rho_L(\theta|\hat{\theta}')f(\theta)d\theta}{F(\hat{\theta})-\Delta} - \frac{\int_{\hat{\theta}'}^{\hat{\theta}} \rho_L(\theta|\hat{\theta}')f(\theta)d\theta}{F(\hat{\theta})-\Delta} > \frac{1-\int_{\overline{\theta}_L}^{\theta_2} \rho_L(\theta|\hat{\theta})f(\theta)d\theta}{F(\hat{\theta})}$. The inequality holds if and only if

$$1 - \frac{\Delta}{F(\hat{\theta})} < \frac{1 - \int_{\hat{\theta}}^{\theta_2} \rho_L(\theta|\,\hat{\theta}')f(\theta)d\theta - \int_{\hat{\theta}'}^{\hat{\theta}} \rho_L(\theta|\,\hat{\theta}')f(\theta)d\theta}{1 - \int_{\hat{\theta}}^{\theta_2} \rho_L(\theta_0|\,\hat{\theta})f(\theta)d\theta}. \tag{19}$$

Because $\rho_L(\theta|\,\hat{\theta}') < \rho_L(\theta|\,\hat{\theta})$, $1 - \int_{\hat{\theta}}^{\theta_2} \rho_L(\theta|\,\hat{\theta}')f(\theta)d\theta > 1 - \int_{\hat{\theta}}^{\theta_2} \rho_L(\theta|\,\hat{\theta})f(\theta)d\theta$. RHS must be greater than $1 - \frac{\int_{\hat{\theta}'}^{\hat{\theta}} \rho_L(\theta|\hat{\theta}')f(\theta)d\theta}{1-\int_{\hat{\theta}}^{\theta_2} \rho_L(\theta_0|\hat{\theta})f(\theta)d\theta}$. Hence, it suffices to show $\frac{\int_{\hat{\theta}'}^{\hat{\theta}} \rho_L(\theta|\hat{\theta}')f(\theta)d\theta}{1-\int_{\hat{\theta}}^{\theta_2} \rho_L(\theta|\hat{\theta})f(\theta)d\theta} < \frac{\Delta}{F(\hat{\theta})}$. This is obvious, because (1) $1 - \int_{\hat{\theta}}^{\theta_2} \rho_L(\theta_0|\,\hat{\theta})f(\theta)d\theta > F(\hat{\theta})$ and (2) $\int_{\hat{\theta}'}^{\hat{\theta}} \rho_L(\theta|\,\hat{\theta}')f(\theta)d\theta < \Delta$. Hence, the payoff from no reform $\alpha \diagup \left[\alpha + (1-\alpha)\frac{1-\int_{\overline{\theta}_L}^{\theta_2} \rho_L(\theta_0|\hat{\theta})f(\theta)d\theta}{F(\hat{\theta})}\right]$ must decrease. Hence, the payoff from no reform must decrease instead of increasing. Contradiction.

**Case 2.3 $\overline{\theta}_L = \hat{\theta}$ and $\overline{\theta}'_L > \hat{\theta}'$.** The proof is similar to that for Case 2.2.

**Case 2.4 $\overline{\theta}_L > \hat{\theta}$ and $\overline{\theta}'_L = \hat{\theta}'$.** The proof is similar to that for Case 2.1.

We conclude that $\rho^*_L(\theta|\,\hat{\theta})$ strictly decreases with $\hat{\theta}$ must decrease with $\hat{\theta}$ for all $\theta \in [\overline{\theta}_L, \theta_2]$.

Further, we establish that $\overline{\theta}_L$ must decrease with $\hat{\theta}$. Again, we prove the claim by contradiction. Assume that $\hat{\theta}$ decreases to $\hat{\theta}'$. Suppose the contrary that $\overline{\theta}_L$ increases to $\overline{\theta}'_L$. Because the low type reforms more for all $\theta \in [\overline{\theta}'_L, \theta_2]$, the low type must have a lower expected payoff for all $\theta$, regardless of undertaking the reform or maintaining status quo. Consider an arbitrary $\theta = \overline{\theta}'_L - \varepsilon > \overline{\theta}_L$, where $\varepsilon$ is an infinitely small positive number. In the new equilibrium, the low type prefers to receive the payoff of no reform for $\theta$, which must be higher than he would deviate. However, the payoff if he deviates must be strictly higher than his payoff in the previous equilibrium, as the public would believe he is the high type with certainty. This contradicts with the fact that the equilibrium payoff in the new equilibrium is lower than that of the previous equilibrium. ∎

## 7.10 Proof of Remark 3

*Proof.* By the Implicit Function Theorem, the derivative of $\rho^*$ with respect to any parameter is equal to the opposite of the ratio between the derivatives of the equilibrium condition (11)

with respect to that parameter and $\rho$, evaluated at $\rho = \rho^*$. For example,

$$\frac{\partial \rho^*}{\partial \alpha} = -\frac{\partial \text{ LHS}-\text{RHS of (11)}/\partial \alpha}{\partial \text{ LHS}-\text{RHS of (11)}/\partial \rho}\bigg|_{\rho=\rho^*}.$$

Now, we make the following observations, which lead to the conclusions in the proposition.

1. $\partial$ LHS$-$RHS of (11)$/\partial\rho > 0$ for all $\rho > 0$.

The left-hand side of (11) is increasing in $\rho$ and the right-hand side is decreasing.

2 $\partial$ LHS$-$RHS of (11)$/\partial\eta < 0$ for all $\rho > 0$.

The left-hand side of (11) is independent of $\eta$, while the right-hand side derivative with respect to $\eta$ is equal to $\frac{1}{2}(\mu^s + \mu^f) - \mu^n$. We claim it is negative. This is basically saying that the low type's expected reputation when the reform outcome is discovered is worse than when it is not. To see this, observe that

$$\begin{aligned}
& \frac{1}{2}\cdot\frac{1}{1+\lambda(\alpha)B} + \frac{1}{2}\cdot\frac{1}{1+\lambda(\alpha)C} - \frac{1}{1+\lambda(\alpha)\rho} \\
= & \frac{[1+\lambda(\alpha)(B+C)/2][1+\lambda(\alpha)\rho] - [1+\lambda(\alpha)B][1+\lambda(\alpha)C]}{[1+\lambda(\alpha)B][1+\lambda(\alpha)C][1+\lambda(\alpha)\rho]}, \\
= & \frac{\lambda(\alpha)[\rho - (B+C)/2] + \lambda(\alpha)^2[\rho(B+C)/2 - BC]}{[1+\lambda(\alpha)B][1+\lambda(\alpha)C][1+\lambda(\alpha)\rho]}, \\
= & \frac{\lambda(\alpha)\rho[1 - \frac{1}{4q(1-q)}]}{[1+\lambda(\alpha)B][1+\lambda(\alpha)C][1+\lambda(\alpha)\rho]}.
\end{aligned}$$

The expression is negative as $4q(1-q) < 1$.

3 $\partial$ LHS$-$RHS of (11)$/\partial\alpha > 0$ at $\rho = \rho^*$.

Note that $\partial$ LHS$-$RHS of (11)$/\partial\alpha = \partial$ LHS$-$RHS of (11)$/\partial\lambda \cdot \lambda'(\alpha)$. As $\lambda'^2 < 0$, it suffices to show that

$$\partial \text{ LHS}-\text{RHS of (11)}/\partial\lambda < 0.$$

Note that

$$\begin{aligned}
& \frac{\partial \text{ LHS}-\text{RHS of (11)}}{\partial\lambda} \\
= & \frac{-A}{[1+\lambda(\alpha)A]^2} + \frac{\eta}{2}\cdot\frac{B}{[1+\lambda(\alpha)B]^2} + \frac{\eta}{2}\cdot\frac{C}{[1+\lambda(\alpha)C]^2} + (1-\eta)\cdot\frac{\rho}{[1+\lambda(\alpha)\rho]^2}, \\
= & \frac{1}{1+\lambda(\alpha)A}\left[\frac{-A}{1+\lambda(\alpha)A} + \frac{\eta}{2}\cdot\frac{1+\lambda(\alpha)A}{1+\lambda(\alpha)B}\cdot\frac{B}{1+\lambda(\alpha)B}\right. \\
& \qquad \left. + \frac{\eta}{2}\cdot\frac{1+\lambda(\alpha)A}{1+\lambda(\alpha)C}\cdot\frac{C}{1+\lambda(\alpha)C} + (1-\eta)\cdot\frac{1+\lambda(\alpha)A}{1+\lambda(\alpha)\rho}\cdot\frac{\rho}{1+\lambda(\alpha)\rho}\right].
\end{aligned}$$

Now observe for any $\rho > 0$, $B < \rho \leq 1 < A$, given their definitions and the fact $q > 3/4$. Now, in order for (11) to be satisfied, it must be the case $C > A$ at $\rho = \rho^*$. We may conclude that, at $\rho = \rho^*$,

$$\begin{aligned} &B < \rho < A < C, \\ &\frac{B}{1+\lambda(\alpha)B} > \frac{\rho}{1+\lambda(\alpha)\rho} > \frac{A}{1+\lambda(\alpha)A} > \frac{C}{1+\lambda(\alpha)C}. \end{aligned}$$

The equilibrium condition (11) implies that

$$\frac{A}{1+\lambda(\alpha)A} = \frac{\eta}{2} \cdot \frac{B}{1+\lambda(\alpha)B} + \frac{\eta}{2} \cdot \frac{C}{1+\lambda(\alpha)C} + (1-\eta) \cdot \frac{\rho}{1+\lambda(\alpha)\rho}.$$

Combining it with the above two inequalities, we can then conclude

$$\frac{\partial \text{ LHS}-\text{RHS of (11)}}{\partial \lambda} < 0$$

at $\rho = \rho^*$.

4 $\partial$ LHS$-$RHS of (11)$/\partial q > 0$ for all $\rho > 0$.

The left-hand side of (11) is independent of $q$, while the right-hand side derivative with respect to $q$ is the same as that for (3), which is negative. ∎

# References


**Alesina, Alberto and Guido Tabellini**, "Bureaucrats or Politicians? Part I: A Single Policy Task," *American Economic Review*, March 2007, *97* (1), 169–179.

**Banks, Jeffrey S. and Joel Sobel**, "Equilibrium Selection in Signaling Games," *Econometrica*, May 1987, *55* (3), 647–61.

**Bierbrauer, Felix and Lydia Mechtenberg**, "Winners and Losers of Early Elections: On the Welfare Implications of Political Blockades and Early Elections," Working Paper Series of the Max Planck Institute for Research on Collective Goods 2008_50, Max Planck Institute for Research on Collective Goods 2008.

**Biglaiser, Gary and Claudio Mezzetti**, "Politicians' decision making with re-election concerns," *Journal of Public Economics*, December 1997, *66* (3), 425–447.

**Brandenburger, Adam and Ben Polak**, "When Managers Cover Their Posteriors: Making the Decisions the Market Wants to See," *RAND Journal of Economics*, Autumn 1996, *27* (3), 523–541.

**Capelos, Tereza**, "Shield or Stinger: The Role of Party Bonds and Competence Evaluations in Political Predicaments," Technical Report, University of Surrey 2005. Paper presented at the annual meeting of the American Political Science Association, available online.

**Caselli, Francesco and Massimo Morelli**, "Bad politicians," *Journal of Public Economics*, March 2004, *88* (3-4), 759–782.

**Chen, Ying**, "Career Concerns, Project Choice, and Signalling," working paper, Arizona State University 2010.

**Chung, Kim-Sau and Péter Esö**, "Signalling with Career Concerns," working paper 2008.

**Dur, Robert A. J.**, "Why Do Policy Makers Stick to Inefficient Decisions?," *Public Choice*, June 2001, *107* (3-4), 221–34.

**Fernandez, Raquel and Dani Rodrik**, "Resistance to Reform: Status Quo Bias in the Presence of Individual-Specific Uncertainty," *American Economic Review*, December 1991, *81* (5), 1146–55.

**Hargrove, Erwin C.**, *Presidential Leadership: Personality and Political Style*, New York: MacMillan Company, 1966.

**Hermalin, Benjamin E.**, "Managerial preferences concerning risky projects," *Journal of Law, Economics, and Organization*, 1993, *9* (1), 127–135.

**Holmström, Bengt**, "Managerial Incentive Problems: A Dynamic Perspective," in B. Wahlroos, B.-G. Ekholm, H. Meyer, and A.-E. Lerviks, eds., *Essays in Economics and Management in Honour of Lars Wahlbeck*, Helsinki: Swedish School of Economics, 1982.

– , "Managerial Incentive Problems: A Dynamic Perspective," *Review of Economic Studies*, January 1999, *66* (1), 169–82.

– **and Joan E. Ricart i Costa**, "Managerial incentives and capital management," *The Quarterly Journal of Economics*, 1986, *101* (4), 835–860.

**Howitt, Peter and Ronald Wintrobe**, "The political economy of inaction," *Journal of Public Economics*, March 1995, *56* (3), 329–353.

**Kwon, Young K.**, "Accounting conservatism and managerial incentives," *Management science*, 2005, *51* (11), 1626–1632.

**Levy, Gilat**, "Decision Making in Committees: Transparency, Reputation, and Voting Rules," *American Economic Review*, March 2007, *97* (1), 150–168.

**Li, Hao**, "A Theory of Conservatism," *Journal of Political Economy*, June 2001, *109* (3), 617–636.

**Li, Wei**, "Changing One's Mind when the Facts Change: Incentives of Experts and the Design of Reporting Protocols," *Review of Economic Studies*, 2007, *74*, 1175–1194.

**Majumdar, Sumon and Sharun W. Mukand**, "Policy Gambles," *American Economic Review*, September 2004, *94* (4), 1207–1222.

**Mattozzi, Andrea and Antonio Merlo**, "Mediocracy," NBER Working Papers 12920, National Bureau of Economic Research, Inc February 2007.

– **and** – , "Political careers or career politicians?," *Journal of Public Economics*, April

2008, *92* (3-4), 597–608.

**Messner, Matthias and Mattias K. Polborn**, "Paying politicians," *Journal of Public Economics*, December 2004, *88* (12), 2423–2445.

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

**Ottaviani, Marco and Peter Norman Sørensen**, "Professional Advice," *Journal of Economic Theory*, 2006, *126* (1), 120–142.

**Prat, Andrea**, "The wrong kind of transparency," *American Economic Review*, June 2005, *95* (3), 862–877.

**Prendergast, Canice and Lars Stole**, "Impetuous Youngsters and Jaded Oldtimers," *Journal of Political Economy*, December 1996, *104* (6), 1105–1134.

**Sanyal, Amal and Kunal Sengupta**, "Cheap Talk, Status Quo and Delegation," SSRN Working Paper Series 2006.

**Scharfstein, David S. and Jeremy C. Stein**, "Herd Behavior and Investment," *American Economic Review*, June 1990, *80* (3), 465–79.

**Sheehan, Frederick**, *Panderer to Power: The Untold Story of How Alan Greenspan Enriched Wall Street and Left a Legacy of Recession*, New York: McGraw-Hill, 2009.

**Skocpol, Theda**, "Bringing the State Back In: Strategies of Analysis in Current Research," in Peter B. Evans, Dietrich Rueschemeyer, and Theda Skocpol, eds., *Bringing the State Back In*, Cambridge: Cambridge University Press, 1985.

**Suurmond, Guido, Otto H. Swank, and Bauke Visser**, "On the bad reputation of reputational concerns," *Journal of Public Economics*, 2004, *88* (12), 2817–2838.

**Tirole, Jean**, "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations," *Journal of Law, Economics, and Organization*, Autumn 1986, *2* (2), 181–214.

**Zwiebel, Jeffrey**, "Corporate Conservatism and Relative Compensation," *Journal of Political Economy*, February 1995, *103* (1), 1–25.