

Politics and Information in Economic Reforms

Zhaotian Luo
Department of Politics
New York University

August 27, 2018

Abstract

Why do economic reforms follow different patterns? Why are some reforms abandoned, some recklessly assume the form of a “big bang,” while some proceed through careful experimentation? A formal model addresses these questions by analyzing political incentives of competing politicians and the effect of these incentives on the information they seek. Three main results are derived. First, economic reforms can entail smaller or larger political stakes and different degrees of politicization can lead to reform gridlock, or result in imprudent big-bang implementation, or encourage careful implementation through experimentation. Second, if reform policies are tested through experiments, the level of politicization affects the information these experiments produce. Third, contingent on their factional alignments and the status quo of economy, politicians have incentives to increase or reduce political stakes as a tactic to advance their interests. Politics matters in the process of economic reforms and it matters through shaping politicians’ incentives to seek information about the efficacy of reform policies.

1. Introduction

Two features are ubiquitous in economic reforms: uncertainty about their economic outcomes and the political stakes entailed in their success or failure. A reform can either improve or damage the functioning of the economy and its effects often remain unknown until it is implemented. Even Deng Xiaoping, the “chief designer” of China’s market economy reform, “did not have a clear blueprint about how to bring wealth to the people and power to the country; instead, as he confessed, repeating a widely used saying — he groped for the stepping stones as he crossed the river” (Vogel, 2011). As for the politics, reforms provide political actors with opportunities to enhance as well as risks to lose political influence. Promoting a reform policy is a political gamble: one gains politically if the policy

turns out to be a success but has to pay if it is a failure. Wan Li, Deng's vanguard, successfully promoted the policy of agricultural de-collectivization (Household Responsibility System Reform) and as a result he replaced Chen Yonggui and Wang Renzhong, who were known for their conservative stance, taking charge of agriculture at the national level (Vogel, 2011). Nikita Khrushchev lost support from within the Party due to his failed policy of agricultural decentralization (Sovnarkhoz Reform), which led to his dismissal in 1964 (Markevich and Zhuravskaya, 2011).

The politics of economic reforms regularly pits "reformers" against "conservatives." The political actors who are in the position to propose policy changes and manage their implementation often form a "reform faction." They may seek, or may be accused by their political rivals of seeking, to grab power under the disguise of reform. Motivated by political opportunism, the reform faction may *over-reform*, promoting radical policy changes without fully understanding their effects, which can lead to hasty implementation of policies that are detrimental to the economy. In turn, political opponents of reforms often form a "conservative faction." Motivated by the fear of losing political power, they may *under-reform*, resisting any attempt to change the status quo, which can lead to abortion of policies that are beneficial to the economy. Therefore, a systematic analysis of political incentives of the relevant decision-makers is necessary to understand why and in what way do some reforms succeed while others fail. The market-economy reform in China under Deng led to a spectacular success while the agriculture reform in the former Soviet Union under Khrushchev ended in utter failure.

In this paper, a model of collective decision-making is developed to study the interactive effects of the economic and political aspects of reforms. The model is based on two patterns of economic reform that are consistent with the two features of reform.

First, *experiments* can be designed and conducted to discover the unknown effects of reform policies. Experimentation can take the form of small scale implementation, local pilots, field surveys, simulations, or debates, etc. — any process that helps to understand the effects of a particular policy on the economy. The method of experimentation, that is, careful decision-making guided by

experiments and the institutional arrangements that facilitated its application, is argued to have played a crucial role in the Chinese reforms (see [Naughton, 1995](#); [Qian, Roland and Xu, 2006](#); [Heilmann, 2008](#)).

Second, reform policies can be *politicized* or de-politicized, escalating or reducing the political stake they entail. The level of politicization can be a result of ideology, institutions, or politicians' strategic choices. In communist Poland, building a public toilet in an agricultural market place could have been politically risky for mayors of small towns, as it would encourage private sale of agricultural products, which was a deviation from the Party Line (see [Luo and Przeworski, 2017](#)). The "Party Line" and the doctrine of complying with it intensely politicized policy changes. In contrast, Deng often resorted to his famous saying "It doesn't matter if the cat is black or white as long as it can catch mice" to diminish the ideological controversy over policies that were widely perceived as capitalist. This move to de-politicize encouraged reform policies such as the de-regulation of private firms ([Vogel, 2011](#)).

To investigate its effect on experimentation, the level of politicization is first treated as given exogenously. This assumption helps to answer questions such as whether reforms would get trapped in a gridlock, whether a particular policy change would be adopted in a big-bang fashion or be subjected to an experiment, and how experiments, if any, would be designed. Only then politicization and de-politicization are considered as maneuvers of politicians and, therefore, as endogenously determined by their strategic interactions. This analysis explains whether and under what conditions would reform policies be politicized as in Poland or de-politicized as in China.

Three main results emerge from the model. First, depending on its level, politicization can lead to reform gridlock or can contribute to imprudence. Specifically, a high level of politicization deters politicians from proposing any policy changes and, therefore, leads to gridlock. But when the status quo of the economy is poor, a positive but relatively low level of politicization encourages big-bang implementation of reform policies and, therefore, results in imprudence.

Second, if a particular reform policy is tested through experimentation, the level of politicization affects the effectiveness of the experiment. Specifically, when the

level of politicization is too low, the experiment would produce *false positives*, yielding supportive evidence about the efficacy of the reform policy when it is in fact detrimental. These false positives could mislead politicians to commit *type-I errors* — to implement reform policies that are harmful to the economy. When the level of politicization is too high, however, the experiment would produce *false negatives*, yielding evidence about the inefficacy of a reform policy when it is in fact beneficial. These false negatives could mislead politicians to commit *type-II errors* — to give up reform policies that improve the economy.

Third, if politicians can use politicization and de-politicization as tactics to advance their interests, they would manipulate the level of politicization contingently on their factional alignments and the status quo of economy. Specifically, politicians in the reform faction would de-politicize when the status quo of economy is satisfactory and would politicize, still to some relatively low level, when the status quo is bad. On the contrary, politicians in the conservative faction would de-politicize when the status quo of economy is fair, but when the status quo is unsatisfactory, they would politicize to a relatively high level.

The rest of the paper is structured as follows. Section 2 presents the formal set-up of the model. Section 3 analyzes how politicians would choose the way to reform under an exogenously given level of politicization. Section 4 investigates politicization as strategic choices of politicians. Section 5 discusses the related literature and concludes the paper.

2. Model Set-Up

The players are a *Reformer* (R), or a reform faction, and a *Conservative* (C), or the opponent faction of R. The game begins with a *status quo* of economy and a level of *politicization* that specifies the political stake of any attempt to reform. R has a choice to *initiate* a reform policy that once carried out, moves the economy from the status quo to a new state. Initiation means bringing a policy change into the political agenda, making it a politically salient issue. Initiation is a necessary yet not sufficient condition for implementing a policy. Given that R has initiated a policy, R and C decide together whether to *implement* it by *consensus*: implementation requires consent of both parties. The status quo is preserved if R

fails to initiate a reform policy or if at least one of the two players vetoes implementation of an initiated policy.

Economic Uncertainty. Economically, R and C share common interests. Both parties derive a payoff of $\mu \in (0, 1)$ from the status quo of the economy and both derive a payoff of $\theta \in [0, 1]$ from implementing a reform policy. Parameters μ and θ measure, respectively, the economic efficacy of the status quo and of the state resulting from the policy change. A policy that improves the economy relative to the status quo, that is, $\theta \geq \mu$, is referred to as an *efficient* reform, while one that damages the economy relative to the status quo, that is, $\theta < \mu$, is *inefficient*.

Considering only the economy, it is in the best interests of both players to carry out an efficient policy change and prevent an inefficient one. The problem, however, is that the economic consequences of a reform policy are often not fully known before its implementation. The two players face economic uncertainty about reform. Formally, the economic efficacy of a reform policy θ is unknown and stochastically drawn from a prior distribution F . Without precise information about θ , implementing a reform policy is an economic gamble for the two players: doing so, they can commit a *Type-I Error*, that is, to implement an inefficient policy. However, by preventing the implementation of a reform, the two players could commit a *Type-II Error*, failing to implement an efficient policy.

Table 1: Errors

	$\theta < \mu$	$\theta \geq \mu$
Status Quo		Type-II Error
Reform	Type-I Error	

Political Stake. Politically, R and C have competing interests. Reform offers both factions political opportunities as well as risks. Once a reform policy is initiated, it becomes politically salient and its implementation or abortion bears political consequences. If R manages to successfully implement a policy change he initiates, he would gain political capital and his advantage against C — a political opponent — would increase. If an initiated policy change is vetoed, R would lose political capital and his advantage against C would diminish. Therefore, the initiation of a reform policy is a political gamble for R: it entails a political stake. Let $b \geq 0$ denote

the scale of this political stake: it measures what R gains and C loses politically when an initiated policy gets implemented, as well as what C gains and R loses when an initiated policy is aborted. Parameter b measures the level of *politicization* on the issue of reforms in the game between R and C. With a larger b , whether or not to implement an initiated policy matters more politically for both parties.

As a summary, R and C have both economic and political incentives. Economically, reform would lead to an uncertain new state of economy that has a chance to be better than the status quo, or worse. Politically, reform can be politicized, creating a political stake in its implementation. The aggregate payoffs of the two players in all the potential scenarios are illustrated in the following table.

Table 2: Payoffs

	R	C
Not Initiated	μ	μ
Initiated, Blocked	$\mu - b$	$\mu + b$
Initiated, Implemented	$\theta + b$	$\theta - b$

Experimentation. The economic uncertainty of reform can be resolved through experimentation. Having initiated a reform policy, R has an option to design and conduct an experiment to test its economic efficacy. Such an experiment may take the form of local pilots, surveys, examinations, or debates etc. — any process that discloses information about the economic efficacy of a particular reform policy.

Formally, an *experiment* is a process that stochastically generates one in two possible results $s = 0, 1$ conditional on the economic efficacy of the initiated policy, where $s = 1$ is a positive result that indicates the experiment succeeded while $s = 0$ is a negative result that indicates the experiment failed. An experiment is fully characterized by a function $q : [0, 1] \rightarrow [0, 1]$ such that $q(\theta) = Pr[s = 1|\theta]$ is the probability that the experiment succeeds conditional on θ .

Experiments can be either *informative* or *uninformative*. For instance, experiment

q^* such that

$$q^*(\theta) = \begin{cases} 0, & \theta < \mu \\ 1, & \theta \geq \mu \end{cases}$$

is informative. The results of this experiment perfectly reveal whether the initiated policy is efficient: each player knows that the reform is efficient once observing $s = 1$ and each knows that it is inefficient once observing $s = 0$. But an experiment that always succeeds is uninformative. It never fails and its success is uninformative about θ .

Timing. The sequence of moves is as follows.

1. The level of politicization b is determined and observed by both players.
2. R decides whether to initiate a reform policy and
 - 2.1 If no policy is initiated, the game ends with the status quo preserved;
 - 2.2 If a policy is initiated, its economic efficacy θ is drawn according to F .
3. R designs and conducts an experiment q .
4. The experiment generates a result $s = 0, 1$ according to $Pr[s = 1|\theta] = q(\theta)$.
5. Both players observe q, s and decide whether to implement the initiated policy.

In the subsequent discussion, b is first assumed to be exogenously given to investigate its effect on the strategic interaction between R and C. Then, b is considered endogenous choices of the two players to examine their incentives to politicize or de-politicize reform. In all cases, the solution concept is *Perfect Bayesian Equilibrium* (equilibrium).

If, in equilibrium, R does not initiate any reform policy, there is a *reform gridlock*. If R initiates a reform policy without experimentation — setting an uninformative experiment, he attempts to reform through a *big-bang approach*. In contrast, if R initiates a reform policy with experimentation — running an informative experiment, he attempts to reform through an *experimentalist approach*.

3. Effects of Politicization

This section studies the effects of politicization on reform — whether would any policy change be initiated and, once initiated, how would reform policies be implemented — for an exogenously given level of politicization.

Obviously, the maximal payoff C can expect from implementing a reform policy is $1 - b$ and by vetoing an initiated policy, he gets $\mu + b$. Hence, if $b > \frac{1-\mu}{2}$, so that $1 - b < \mu + b$, C would veto any policy change R initiates and in turn, R would never initiate one. Therefore, without loss of generality, the subsequent analysis focuses on the non-trivial case in which $b \leq \frac{1-\mu}{2}$.

3.1. Experimentation

Suppose R has initiated a reform policy, making it politically salient, then he has an option to conduct an experiment to test the economic efficacy of the initiated policy. Note that R can as well try to implement the initiated policy through a big-bang approach by setting an uninformative experiment.

Lemma 1. *If R initiates a reform policy, he would set an experiment q such that*

$$q(\theta) = \begin{cases} 0, & \theta < k \\ 1, & \theta \geq k \end{cases}.$$

Under this experiment, the initiated policy would be implemented if the experiment succeeds and vetoed if the experiment fails.

The above lemma has two results. The second result shows that the collective decision of the two players with regard to the implementation of the reform policy R initiates is completely determined by the experimental result they observe. Both parties would approve if the experiment shows a positive result and at least one player would disapprove if the experiment draws a negative result. Hence, experimental results theoretically can be categorized into four groups. First, $s = 1$ is a *True Positive* (TP) when $\theta \geq \mu$, in which case the positive result correctly points out that the initiated policy is efficient and guides the players to implement it. But

$s = 1$ is a *False Positive* (FP) when $\theta < \mu$, in which case it misleads the players to make a type-I error by carrying out an inefficient policy change. Similarly, $s = 0$ is a *True Negative* (TN) when $\theta < \mu$, in which case the negative result correctly warns about the inefficiency of the initiated policy and guides the players to abort it. But $s = 0$ is a *False Negative* (FN) when $\theta \geq \mu$, in which case the experimental failure misleads the players to commit a type-II error by missing a good opportunity to improve the economy.

Table 3: Experimental Results

	$\theta < \mu$	$\theta \geq \mu$
Failure	True Negative	False Negative
Success	False Positive	True Positive

The first result of Lemma 1 indicates that R would always choose an experiment that is fully characterized by a threshold value k . Such an experiment would succeed if and only if the initiated reform policy has a large enough economic efficacy that exceeds the threshold, that is, if and only if $\theta \geq k$. This threshold measures how difficult it is for the experiment to succeed. It has both economic and political implications.

To see the economic implication, consider first q^* defined in the set-up. This experiment has an intermediate threshold μ . It succeeds if and only if the initiated reform policy is efficient and does not draw any false positive or false negative. If R sets this experiment, he and C would implement the initiated policy if it is economically superior to the status quo and block it if it is inferior: their collective decision would always be correct. For this reason, q^* is referred to as the *first-best* experiment. Any deviation from the first best can lead to erroneous joint decisions.

On the one hand, if R sets an experiment that is too easy to produce a positive result, that is, if $k < \mu$, the two players face the possibility of committing type-I errors. Specifically, if the initiated reform policy is moderately inefficient, so that $k \leq \theta < \mu$, the experiment would generate a false positive that misleads the two players to implement an economically detrimental policy change. Interestingly, the experiment in this case would never produce a false negative. Because $k < \mu$, the experiment would fail only if the initiated policy is inefficient. On the other hand, if

R sets an experiment that is too difficult to generate a positive result, that is, if $k > \mu$, the two players may face the possibility of committing type-II errors. Specifically, if the initiated policy is efficient but only has an intermediate efficacy, so that $\mu \leq \theta < k$, the experiment could produce a false negative that misleads the two players to abort an economically improving policy change. Similarly to the previous case, the experiment in this case would never produce a false positive. Because $k > \mu$, the experiment would succeed only if the initiated policy is efficient.

The threshold of the experiment R sets also affects probability for him to gain political advantage against C and to avoid a political loss. Given that the level of politicization is b , R would enjoy the political benefit of b by successfully implementing the initiated reform policy and would suffer the political loss of b by failing to do so. The probability of implementation — the probability for R to get the political benefit — is $1 - F(k)$, which is strictly decreasing in the threshold k . Hence, by making the experiment easier to succeed, R gets a larger chance to gain politically at the expense of C.

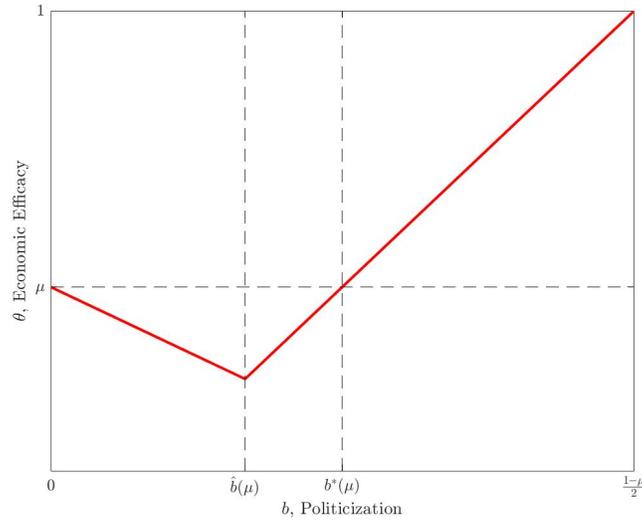
To summarize the above discussion, economically, R would want to choose an intermediate threshold for his experiment, that is, one as close as possible to μ , to guide himself and C to make economically wise joint decisions. Politically, R would want to choose a threshold that is as low as possible to have a good chance to implement the initiated reform policy and realize the political benefit it entails.

Proposition 1. *If R initiates a reform, he has a unique optimal experiment which is characterized by the threshold value*

$$k^*(b, \mu) := \min \{k \geq \mu - 2b : \mathbb{E}[\theta | \theta \geq k] \geq (\mu + 2b)\}.$$

Moreover, there exists a unique $\hat{b}(\mu) \in (0, \frac{1-\mu}{2})$ such that $k^*(b, \mu) = \mu - 2b$ if $b \leq \hat{b}(\mu)$ and is increasing in b if $b > \hat{b}(\mu)$.

This proposition characterizes R's optimal experimental design and, most importantly, specifies how it depends on the level of politicization. As illustrated in Figure 1, when the level of politicization b increases, the threshold of R's optimal experiment $k^*(b, \mu)$ first decreases until b reaches a critical value $\hat{b}(\mu)$ and then increases. To understand the effect of politicization, remember that implementation



Red line is $k^*(b, \mu)$
 $F(x) = x$ and $\mu = 0.4$

Figure 1: R's Optimal Experiment

of the initiated reform policy requires consent by both R and C and the level of politicization affects the incentives of both parties.

First, as a benchmark for subsequent discussion, suppose that reform is a completely depoliticized issue, that is, $b = 0$. In this case, R is concerned only with the economy. Hence, he would want to set the first-best experiment q^* . Under this experiment, if C observes an experimental success, he expects to achieve a positive net benefit of $\mathbb{E}[\theta|\theta \geq \mu] - \mu$ by implementing the reform policy and would strictly prefer to do so. As a result, q^* is optimal in this case and $k^*(0, \mu) = \mu$.

Now suppose reform is a somewhat politicized issue, that is, $b > 0$. A positive level of politicization has two effects. On the one hand, due to the positive political stake, R's payoff would increase if the initiated reform policy gets implemented and would decrease if it is blocked. Hence, a positive level of politicization incentivizes R to set an experiment that has a low threshold, making it easier to draw positive results, even at the cost of implementing an economically inefficient reform. Formally, the (ex-post) net payoff R gets by implementing the initiated policy change is $\theta + b - (\mu - b) = \theta - (\mu - 2b)$. Hence, R finds it profitable to carry out the initiated policy as long as $\theta \geq \mu - 2b$. He would want to set an experiment that has

the threshold of $\mu - 2b$. This threshold is strictly decreasing in b . On the other hand, the positive political stake would decrease C's payoff from implementing the initiated reform and increase his payoff from vetoing it. Hence, a positive level of politicization incentivizes C to block the initiated policy change. In turn, to persuade C to approve the initiated policy change, the positive result of R's experiment has to offer a stronger evidence for the economic efficacy of the policy change. Formally, once observing a positive result, C would prefer to implement the initiated policy if and only if his net payoff by doing so, $\mathbb{E}[\theta - b | \theta \geq k] - (\mu + b) = \mathbb{E}[\theta | \theta \geq k] - (\mu + 2b)$, is positive. This net payoff is strictly increasing in k and strictly decreasing in b .

Therefore, a positive level of politicization has a negative effect on the threshold R chooses for his experiment through R's incentive while it has a positive effect through C's incentive. Suppose b is sufficiently small, so that $\mathbb{E}[\theta | \theta \geq \mu - 2b] > \mu + 2b$ or, equivalently, $b < \hat{b}(\mu)$. Then, if R sets an experiment that has his favorite threshold $\mu - 2b$, the positive result this experiment generates would successfully persuade C to approve the implementation of the initiated reform. In this case, b affects $k^*(b, \mu)$ through R's incentive. A marginal increase in b would encourage R's political opportunism, incentivizing him to choose a smaller threshold for his experiment. It follows that $k^*(b, \mu) = \mu - 2b$, which is strictly decreasing in b if b is sufficiently small.

Now suppose b is sufficiently large, so that $\mathbb{E}[\theta | \theta \geq \mu - 2b] \leq \mu + 2b$ or, equivalently, $b \geq \hat{b}(\mu)$. Then, R can no longer choose his favorite threshold $\mu - 2b$ for his experiment. Due to the high political stake, the positive result such an experiment produces is too weak to induce C's support for implementation. Instead, R would set a threshold that is minimally necessary to persuade C. In this case, b affects $k^*(b, \mu)$ through C's incentive. A marginal increase in b would increase C's political loss of implementing the initiated reform policy and in turn, would force R to choose a larger threshold for his experiment.

Corollary 1. *The following hold for $k^*(b, \mu)$,*

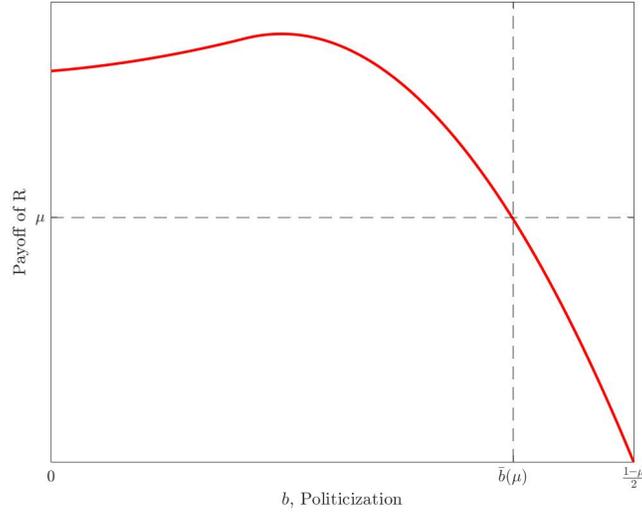
1. $k^*(b, \mu) = \mu$ if and only if $b = 0$ or $b = b^*(\mu) := \frac{\mathbb{E}[\theta | \theta \geq \mu] - \mu}{2} > \hat{b}(\mu)$;
2. $k^*(b, \mu) < \mu$ if $0 < b < b^*(\mu)$ and $k^*(b, \mu) > \mu$ if $b^*(\mu) < b \leq \frac{1-\mu}{2}$.

This corollary shows some important properties of $k^*(b, \mu)$. First, for two particular levels of politicization, $b = 0$ and $b = b^*(\mu)$, R would choose the first-best experiment. Second, if the level of politicization is positive but sufficiently low, so that $0 < b < b^*(\mu)$, R would set an experiment that never produces false negatives but sometimes draws false positives. In this case, the two players reform too much and so they face the danger of implementing a reform policy that damages the economy. If the level of politicization is sufficiently high, so that $b > b^*(\mu)$, R would set an experiment that never produces false positives but sometimes draws false negatives. In this case, the two players reform too little, potentially missing an opportunity to improve the economy by implementing an efficient policy change.

In particular, note that when $0 < b < b^*(\mu)$, the probability for the two players to commit a type-I error, which is $F(\mu) - F(k^*(b, \mu))$, is non-monotone in b . It first increases with b , reaching its maximum at $b = \hat{b}(\mu)$, and then decreases with b and converges to 0 when b approaches $b^*(\mu)$. As mentioned above, when $b < \hat{b}(\mu)$, politicization affects R's experimental design through his own incentive. In this case, as the level of politicization increases, R is encouraged to allow more false positives in his experiment. When $\hat{b}(\mu) < b < b^*(\mu)$, politicization affects R's experimental design through C's incentive. In this case, as the level of politicization increases, R is forced to get rid of some false positives in his experiment. Here, politicization serves as an institutional barrier to regulate R's political opportunism and therefore helping the two players to avoid type-I errors.

3.2. Reform Gridlock and "Big Bang"

The analysis of the last section presumes that R has initiated a reform policy. The decision whether to initiate a policy, however, is endogenous to the level of politicization. As long as reform is a politicized issue, initiating a policy change is a political gamble. Reformers accumulate political capital when the initiatives they promote become general policies. Their competence is verified. More importantly, they get the authority to arrange the specificities in the implementation of the reform policies they initiate, from which they can benefit themselves or their followers. However, reformers lose political capital when the policies they support are aborted. Their capability is questioned. Even worse, political opponents of the reformers would take the aborted policies as opportunities to demote them as well



Red line is $u_R^*(b, \mu)$
 $F(x) = x$ and $\mu = 0.4$

Figure 2: R's Decision on Initiation

as their followers. Hence, politicization must affect R's decision regarding initiation. The question is how would the opportunity politicization provides encourage R or would the risk it imposes on R discourage him?

Proposition 2. *There exists a unique $\bar{b}(\mu) \in (0, \frac{1-\mu}{2})$ such that R would initiate a reform policy if $b \leq \bar{b}(\mu)$ and he would not initiate any policy if $b > \bar{b}(\mu)$. Moreover, if $\mu \leq \frac{\mathbb{E}[\theta]}{2}$ and $\hat{b}(\mu) \leq b \leq \frac{\mathbb{E}[\theta]-\mu}{2}$, $k^*(b, \mu) = 0$ and R would initiate a reform policy and implement it through the big-bang approach.*

The first result of this proposition verifies a conventional wisdom: over-politicization forestalls any attempt to reform the status quo. This result provides an explanation for reform stalemates. Remember that R and C have a common goal to improve the economy. Hence, the political rivalry or competition between the two parties could lead to reform stalemate that in general would harm the realization of their common goal if the political stake is sufficiently large.

To understand this result, note that R has two concerns when deciding whether to initiate a reform policy. One is the economic consequence of the initiated policy change: whether would it improve the economy if implemented. As shown in the

previous section, this concern can be resolved through experimentation. The other concern is the political risk: if the policy R initiates is aborted, he has to pay a political cost. Formally, the expected payoff R achieves by initiating a reform policy is

$$u_R^*(b, \mu) := (1 - F(k^*(b, \mu))) \underbrace{(\mathbb{E}[\theta|\theta \geq k^*(b, \mu)] + b)}_{\text{Implementation}} + F(k^*(b, \mu)) \underbrace{(\mu - b)}_{\text{Abortion}}.$$

This expected payoff consists of two parts: what R gets when the initiated policy gets implemented, that is, $\mathbb{E}[\theta|\theta \geq k^*(b, \mu)] + b$, and what he gets when the initiated policy gets aborted, that is, $\mu - b$. If R has a miserable probability to get the policy he initiates implemented, that is, if $1 - F(k^*(b, \mu))$ is very close to 0, as $\mu - b < \mu$, R would rather not reform at all.

Consider first the benchmark of $b = 0$. In this case, reform is a completely depoliticized issue, so that R is concerned only with improving the economy. Given that R has initiated a reform policy, he can run the first-best experiment that perfectly reveals whether or not the initiated reform is efficient. If the experiment succeeds, R gains the net benefit from implementing an economically efficient policy change. If the experiment fails, R loses nothing because reform is depoliticized. Therefore, as illustrated in Figure 2, R always prefers to initiate. Formally, the expected payoff R gains by initiating a reform policy is

$$u_R^*(0, \mu) = \mu + (1 - F(\mu)) (\mathbb{E}[\theta|\theta \geq \mu] - \mu),$$

which is always greater than his payoff μ by not initiating any reform policy.

Now consider a positive level of politicization. When reform is a politicized issue, R cannot ignore the political risk reform entails. R knows that it is possible that the initiated reform policy may fail to be carried out and if this happens, he has to suffer a political loss of b . Moreover, because implementation also requires consent of C, R has to consider the incentive of C. Politicization provides C with an opportunity to gain political advantage against R by rejecting the policy R initiates. Hence, to get C's support, R has to set an experiment that is difficult to succeed, whose positive results are very informative about high economic efficacy of the initiated reform. But such an experiment has a low probability to draw a positive result and so the probability that the initiated policy gets implemented is low. Hence, when reform is politicized and the political stake is high, R would be reluctant to initiate any

reform policy. Doing so, he does not have a good chance to successfully carry out the initiated policy and when he fails, the political loss is considerable. Formally, the expected payoff of R by initiating a reform policy is

$$u_R^*(b, \mu) = \mu - b + (1 - F(k^*(b, \mu))) (\mathbb{E}[\theta | \theta \geq k^*(b, \mu)] - (\mu - 2b)).$$

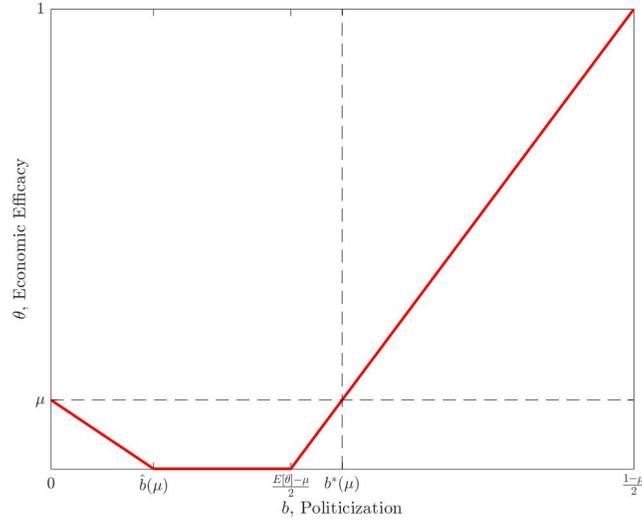
By Proposition 1, $k^*(b, \mu)$ is strictly increasing in b if b is sufficiently large and converges to 1 as b approaches $\bar{b}(\mu)$. Hence, the probability for R to get any policy he initiates implemented, $1 - F(k^*(b, \mu))$, is approximately 0 when the level of politicization is overly high. As a result, R would achieve the expected payoff of approximately $\mu - b$ by initiating a reform policy: he would rather stay with the status quo while not trying to reform at all.

The second result of Proposition 2 shows the conditions under which reform would proceed in the form of a “big bang.” Specifically, when the status quo of economy is poor, so that $\mu \leq \frac{\mathbb{E}[\theta]}{2}$, and when the level of politicization is intermediate, so that $\hat{b}(\mu) \leq b \leq \frac{\mathbb{E}[\theta] - \mu}{2}$, R would initiate a policy change and get it implemented without experimentation. To understand the intuition behind this result, note that economically, a big-bang implementation can potentially hurt both R and C. They both want to avoid making a type-I error by carrying out a policy that harms the economy. When the status quo of economy is poor and the level of politicization is not too high, C would approve a policy change without experimentation. In this case, the status quo is unbearable while the political loss C suffers from reform is acceptable. When the level of politicization is not too low, R would have an incentive to reform without experimentation as the political benefit he obtains from doing so outweighs the expected loss from type-I errors.

3.3. Welfare

Politics, in the model, is a zero-sum game between R and C. Through distorting the incentives of the two players, however, politics can affect the social welfare in general. Specifically, *social welfare* is defined as

$$w^*(b, \mu) := (1 - F(k^*(b, \mu))) \mathbb{E}[\theta | \theta \geq k^*(b, \mu)] + F(k^*(b, \mu)) \mu.$$



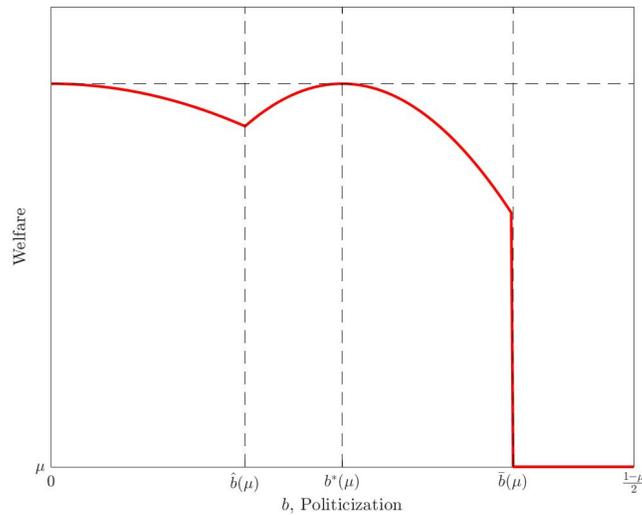
Red line is $k^*(b, \mu)$
 $F(x) = x$ and $\mu = 0.15$

Figure 3: Big-Bang Implementation

There are two interpretations of $w^*(b, \mu)$. First, it is the average of the payoffs of R and C. Second, for people in the society who are “policy-takers” and do not care about elite politics, $w^*(b, \mu)$ measures their expected payoff from policies made by R and C. Note that the level of politicization affects the social welfare only through its effect on R’s experimental design.

Corollary 2. *If $\mu > F^{-1}(\frac{3}{4})$, $w^*(b, \mu)$ is maximized at $b = 0$. If $\mu \leq F^{-1}(\frac{3}{4})$, $w^*(b, \mu)$ is maximized at $b = 0$ and $b = b^*(\mu)$, strictly decreasing in b if $0 < b < \hat{b}(\mu)$, and strictly increasing in b if $\hat{b}(\mu) < b < b^*(\mu)$.*

The above corollary shows how the level of politicization affects social welfare. The effect of politicization is contingent on the economic efficacy of the status quo. First, if the status quo of economy is sufficiently good, that is, if $\mu > F^{-1}(\frac{3}{4})$, it is best to have reform as a completely depoliticized issue. In this case, the conventional wisdom that politicization always hinders social welfare would hold. If the status quo is sufficiently bad, so that $\mu \leq F^{-1}(\frac{3}{4})$, social welfare is maximized not only when reform is completely depoliticized but also when reform is moderately politicized at the level of $b^*(\mu)$. Moreover, if the level of politicization is positive and relatively low, so that $\hat{b}(\mu) < b < b^*(\mu)$, slightly increasing the level of



Red line is $w^*(b, \mu)$
 $F(x) = x$ and $\mu = 0.4$

Figure 4: Social Welfare

politicization would improve the social welfare.

This result is due to the different effects of politicization on the incentives of R and those of C. As mentioned in the previous section, politicization encourages R’s political opportunism. If reform is politicized, R would have a stronger incentive to implement any reform policy he initiates even at the cost of damaging the economy. However, politicization also encourages C’s political conservatism. If R can accumulate the political advantage against C by successfully implementing reform policies, C would have a stronger incentive to veto any policy R initiates. When the level of politicization is positive and relatively low, the political conservatism of C serves as an institutional barrier to regulate R’s political opportunism, preventing R to manipulate his experiment to facilitate the implementation of an inefficient policy change solely due to political incentives.

4. Rise of Politicization

In the previous discussion, the level of politicization is assumed to be exogenous. Nevertheless, in real politics, politicization as well as depoliticization are often

By the above proposition, R does not have a constant preference over the level of politicization. Instead, his preference is contingent on the economic status quo. Specifically, R prefers to completely depoliticize the issue of reform when the status quo of economy is sufficiently good, so that $\mu \geq \mu_R^*$. When the status quo of economy is poor enough, so that $\mu < \mu_R^*$, R would like to have a moderate level of politicization.

For R, politicization is a double-edged sword. It would increase R's payoff when he successfully carries out the reform he initiates, in which case R can exploit the chance of reform to advance his political position. However, politicization would decrease R's payoff when he fails to implement the reform policy he promotes, in which case his failure gives C an excuse to reduce R's political influence. It follows that whether would R prefer to politicize or depoliticize the issue of reform depends on the probability for him to get an initiated policy change implemented.

When the status quo of economy is already quite satisfying, neither R nor C is eager to reform, as the likelihood that any policy change would improve relative to the status quo is low. As a result, the chance for R to get any reform policy he initiates accepted is not promising. In this case, politicization more likely would harm R. Hence, R would find it profitable to depoliticize. When the status quo is economically depressing, the two players are disposed to change, as the chance that any new state would be better than the status quo is large. Consequently, the probability for R to carry out any policy he promotes is promising. In this case, politicization more likely would help R. Hence, R would be better off to politicize.

In the case that R prefers to politicize, however, R cannot overdo it. By Corollary 1, R would always set an experiment that sometimes draws false positives when the level of politicization is positive but sufficiently small, so that $0 < b < b^*(\mu)$. Then, the result in Proposition 3 that $b_R^*(\mu) < b^*(\mu)$ implies that if R finds it profitable to politicize, he would choose a level of politicization under which his optimal experimental design permits false positives. By having false positives, although R may sometimes carry out an inefficient reform policy from which he suffers economically, he increases the chance to implement the policy he initiates and so to further political interests.

To summarize, if R has the option to choose the level of politicization, he would

depoliticize when the status quo of economy is good and as a result, the two players would always make economically correct decisions. R would politicize moderately when the status quo of economy is bad and consequently, the two players have a chance to make type-I errors that damage the economy.

4.2. *C's Optimal Level of Politicization*

Now suppose C has the option to set the level of politicization before R chooses whether to initiate any reform policy. As the political opponent of R, C may want to politicize the issue of reform to regulate R's political opportunism under the disguise of reform and let R pay for "politically incorrect" initiations. Doing so, however, would expose C himself to a political risk: if R manages to get a reform policy he initiates implemented, C would lose political influence to R.

C has to consider two factors when deciding whether and to what extent to politicize the issue of reform. First, by setting an overly high level of politicization $b > \bar{b}(\mu)$, C would deter any attempt of R to reform. In this case, the status quo is always preserved from which C derives a payoff of μ . But the implementation of any reform policy R initiates requires C's consent and C can always secure a payoff of μ by rejecting. Hence, C would always choose a proper level of politicization that allows R to initiate reform policies.

Second, the level of politicization C sets would affect R's experimental design to test the reform policy he initiates and so affects the probability that the initiated policy gets implemented. Specifically, if C chooses $b \leq \bar{b}(\mu)$, R would initiate a policy and run an experiment characterized by $k^*(b, \mu)$ to test its efficacy. In turn, with probability $1 - F(k^*(b, \mu))$, the experiment would succeed, the initiated policy would be implemented, and C would get the expected payoff of $\mathbb{E}[\theta | \theta \geq k^*(b, \mu)] - b$; and with probability $F(k^*(b, \mu))$, the experiment would fail, the initiated policy would be blocked, and C would get the payoff of $\mu + b$. Hence, C achieves the expected payoff of

$$u_C^*(b, \mu) := (1 - F(k^*(b, \mu))) (\mathbb{E}[\theta | \theta \geq k^*(b, \mu)] - b) + F(k^*(b, \mu)) (\mu + b)$$

by choosing some $b \leq \bar{b}(\mu)$.

reform, that is, by setting $b = 0$, C may enjoy the economic improvement of an efficient policy change without suffering any political loss. By choosing the maximal level of politicization $\bar{b}(\mu)$ that allows initiation, C forces R to choose an experiment that has a high threshold to succeed, which ensures that the probability to have a positive experimental result and the policy R initiates gets implemented is low, so that C has a good chance to gain politically from R's failed initiation. In turn, in the case that R successfully carries out the reform policy he initiates, the high threshold value for his experiment guarantees that the reform policy has a very high economic efficacy.

When the status quo of economy is very poor, both C and R are eager to reform, as the likelihood that any policy change would make things worse is quite low. As a result, R has a considerable chance to get a reform policy he initiates accepted. In this case, politicizing the issue of reform would most likely harm C by giving R an opportunity to accumulate political capital. Hence, it is sensible for C to focus on the economic aspect of the reform, to benefit from improvement of the economy resulting from efficient policy changes.

When the status quo of economy is already satisfying, the two players do not have much to gain economically from reform. Hence, in this case, it is more profitable for C to focus on the political aspect of reform. To do so, C can set the level of politicization $\bar{b}(\mu)$ to maximize his political gain from R's failed attempt to promote a policy change.

5. Conclusion

Why do economic reforms follow different patterns? Why are some reforms aborted, some recklessly assume the form of a "big bang," while some proceed through careful experimentation? The model presented here addresses these questions by analyzing political incentives of the relevant decision-makers and the effect of incentives on the information they seek. Specifically, the contribution of this analysis lies in characterizing the conditions under which competing political incentives result in economically beneficial reforms being abandoned or economically deleterious reforms being carried out. Moreover, it is shown that political incentives affect reform strategies by distorting politicians' strategic

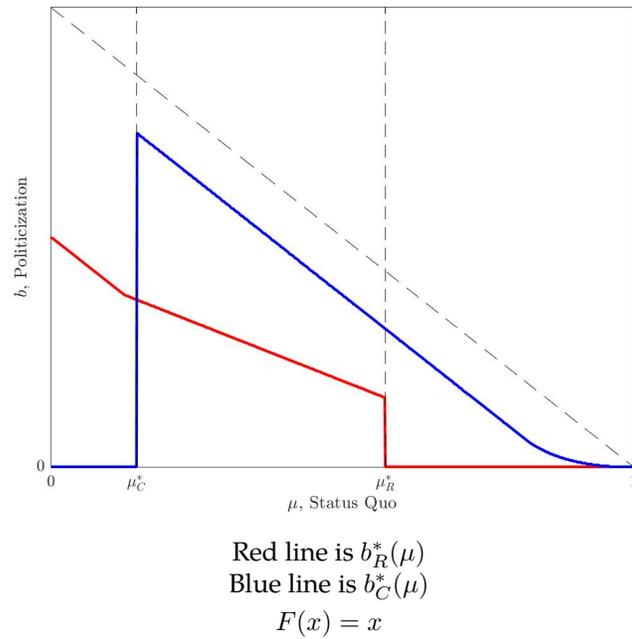


Figure 7: Players' Optimal Level of Politicization

choices with regard to the kinds of information they seek to generate.

The literature on economic reforms, particularly with the focus on the contrast between China, the Soviet Union, and post-communist countries in Eastern Europe, is enormous. Several papers draw comparisons between the big-bang (or "shock therapy") approach and eclectic approaches such as gradualism (McMillan and Naughton, 1992; Dewatripont and Roland, 1992, 1995; Wei, 1997) and dual-track approach (or partial reform) (Murphy, Shleifer and Vishny, 1992; Brandt and Zhu, 2000; Lau, Qian and Roland, 2000; Che and Facchini, 2007; Lin and Li, 2008). These papers emphasize the asymmetric distribution of costs and benefits of reforms. In this light of these papers, reform are feasible only if the losers, those who would bear the costs, can be appeased. Gradualism or dual-track approach can help appease the losers, neutralize their opposition, and therefore facilitate reforms. Hence, reforms that take the "correct" eclectic approach are likely to succeed while those occur in the wrong way are likely to fail. For instance, Wei (1997) argues that gradual implementation of reform policies may help reformers to exploit the divergence of interests among conservatives to neutralize their oppositions; while Lau, Qian and Roland (2000) shows that the dual-track approach is a smart innovation that made the price reform in China one "without losers."

In terms of the emphasis on politics, the current paper is closest to Xie and Xie (2017), where decision-makers have “mutually inclusive payoffs” with regard to the economy but “mutually exclusive payoffs” in politics. They argue that when politicians disagree about the potential effects of reform, that is, when reformers believe that reform is likely to improve the economy while conservatives believe that it is likely to harm it, political incentives may help to break reform stalemate by encouraging them to agree on experimenting with reform policies. The model here, however, does not build on heterogeneous prior beliefs. Learning about reform policies is treated as endogenous, subject to political incentives of both reformers and conservatives. As Shirk (1993) observes, during the reform era of China, some experiments that were supposed to test efficacy of certain reform policies were more or less manipulated, so that they could hardly fail even if these policies would have been in fact detrimental to the economy. While Tu (2008) and Li (2010) mention that some experiments were deliberately made too difficult to pass, so that they could hardly succeed even if the tested policies would improve the economy, the model here assumes that politicians are able to choose how to experiment with reform policies and thus provides a novel explanation for how they would manipulate information experiments reveal.

The key to understanding the patterns of economic reforms lies in the fact that these reforms may or may not be politicized, affecting the political stakes of rival decision-makers. In particular, high political stakes make reforms too risky as a political move, generating stalemates. In turn, a relatively low but positive level of politicization could make reform too attractive a political opportunity, contributing to the big-bang approach. In between the two extreme cases, reform would proceed through experimentation. The informativeness of experimentation, in turn, is also affected by the extent to which reform is politicized. Experiments are more likely to mislead politicians to over-reform when the level of politicization is low while they are more likely to mislead politicians to under-reform when the level of politicization is high.

Furthermore, politicians have incentives to politicize or de-politicize the issue of reform and if they have the option, their strategic choices are contingent on the status quo of economy. To avoid the political risk entailed in policy changes, politicians in the reform faction would prefer to de-politicize reforms when the

status quo of economy is fair, while to exploit the political gains from policy changes, they would prefer to politicize, albeit to a moderate level, when the status quo is bad. Politicians in the conservative faction would opt to de-politicize reforms when the status quo of economy is poor, to avoid losing political power to the reform faction, and they would politicize to a relatively high level when the status quo is good, to benefit politically from failures of the reform faction in promoting policy changes. Note that the status quo of economy in China was poor in 1978, especially in agriculture, before the market-economy reform while reformers were in power, due to several failed policy attempts by Hua Guofeng — a conservative leader (Shirk, 1993). Consistent with the model's prediction, Wan Li, a reformer in agriculture, politicized the Household Responsibility System reform and extended reformers' political influence against conservatives by extending the implementation of this reform to the entire county (Vogel, 2011). In contrast, the status quo of economy was not as bad in the later 1980s, before the price reform. As a result, the price reform was de-politicized and according to McMillan and Naughton (1992), political salience of this reform was quite low — it was done “not by grand policy, but by stealth.”

Competition for political power is crucial for the success and failure of economic reforms. Because any policy change was highly politicized in the former Soviet Union under Brezhnev, adjusting its poorly functioning system of planned economy was too dangerous a political move, and as a result, economic reforms were abandoned. In contrast, because the issue of reforms was de-politicized in China during late 1980s, some reform policies, especially price deregulation, were implemented without being carefully tested, resulting in undesirable consequences (McMillan and Naughton, 1992; Naughton, 1995). Politics matters in the process of economic reform and it matters through shaping politicians' incentives to seek information about the efficacy of reform policies.

References

- Brandt, Loren and Xiaobo Zhu**, "Redistribution in a decentralized economic growth and inflation in China under reform," *Journal of Political Economy*, 2000, 108 (2), 422–439.
- Che, Jiahua and Giovanni Facchini**, "Dual track reforms: with and without losers," *Journal of Public Economics*, 2007, 91, 2291–2306.
- Dewatripont, Mathias and Gerard Roland**, "Economic reform and dynamic political constraints," *Review of Economic Studies*, 1992, 59, 703–730.
- **and Gérard Roland**, "The design of reform packages under uncertainty," *American Economic Review*, 1207-1223 1995, 85 (5).
- Heilmann, Sebastian**, "From local experiments to national polity: the origins of China's distinctive policy process," *The China Journal*, 2008, 59, 1–30.
- Lau, Lawrence, Yingyi Qian, and Gérard Roland**, "Reform without losers: an interpretation of China's dual-track approach to transition," *Journal of Political Economy*, 2000, 108 (1), 120–143.
- Li, Lanqing**, *Breaking through: the birth of China's opening-up policy*, Oxford University Press, 2010.
- Lin, Yifu and Zhiyun Li**, "Policy burden, privatization, and soft budget constraint," *Journal of Comparative Economics*, 2008, 36, 90–102.
- Luo, Zhaotian and Adam Przeworski**, "One-Party State," *Working Paper*, 2017.
- Markevich, Andrei and Ekaterina Zhuravskaya**, "M-form hierarchy with poorly-diversified divisions: a case of Khrushchev's reform in Soviet Russia," *Journal of Public Economics*, 2011, 95, 1550–1560.
- McMillan, John and Barry Naughton**, "How to reform a planned economy: lessons from China," *Oxford review of economic policy*, 1992, 8 (1), 130–143.
- Murphy, Kevin, Andrei Shleifer, and Robert Vishny**, "The transition to a market economy: pitfalls of partial reform," *Quarterly Journal of Economics*, 1992, 107 (3), 889–906.

- Naughton, Barry**, *Growing out of the plan: Chinese economic reform, 1978-1993*, Cambridge University Press, 1995.
- Qian, Yingyi, Gérard Roland, and Chenggang Xu**, "Coordination and experimentation in M-form and U-form organizations," *Journal of Political Economy*, 2006, 114 (2), 366–402.
- Shirk, Susan**, *The political logic of economic reform in China*, University of California Press, 1993.
- Tu, Qiao**, *Biography of Yuan Geng: site of reform and opening-up, 1978–1984*, Writers Publishing House, 2008.
- Vogel, Ezra**, *Deng Xiaoping and the transformation of China*, Belknap Press of Harvard University Press, 2011.
- Wei, Shangjin**, "Gradualism versus big bang: speed and sustainability of reforms," *Canadian Journal of Economics*, 1997, 30 (4b), 1234–1247.
- Xie, Yinxin and Yang Xie**, "Machiavellian experimentation," *Journal of Comparative Economics*, 2017, 45 (4), 685–711.

Appendix

Lemma 1

Proof of Lemma 1. R's ex-post payoff from carrying out the initiated policy is $\theta + b$ and that from vetoing it is $\mu - b$ while C's ex-post payoff from carrying out the initiated policy is $\theta - b$ and that from vetoing it is $\mu + b$. Hence, R would approve implementation as long as C does so. Note that it is without loss of generality to focus on experiments such that $\mathbb{E}_q[\theta|s = 1] \geq \mathbb{E}_q[\theta|s = 0]$.

First, suppose R sets an experiment q under which C would veto when observing $s = 1$. Then, C would also veto when $s = 0$. But by setting q^* , R can get a higher expected payoff. Now suppose R sets an experiment q under which C would approve when $s = 0$. Then, C would also approve when $s = 1$. In turn, R can ensure the same expected payoff by setting an experiment that always succeeds. Therefore, in equilibrium, R always sets an experiment under which C would approve if and only if $s = 1$.

Suppose R sets q . Then, C prefers to implement the reform after observing $s = 1$ if and only if

$$\int_0^1 (\theta - \mu - 2b)q(\theta)f(\theta)d\theta \geq 0.$$

Assume that C prefers to implement the reform if and only if he observes $s = 1$, R's expected payoff is

$$\begin{aligned} \mu + \int_0^1 ((\theta - \mu + b)q(\theta) - b(1 - q(\theta)))f(\theta)d\theta = \\ \mu - b + \int_0^1 (\theta - \mu + 2b)q(\theta)f(\theta)d\theta. \end{aligned}$$

Therefore, any optimal experiment of R solves

$$\begin{aligned} \max_q \int_0^1 (\theta - \mu + 2b)q(\theta)f(\theta)d\theta \\ \text{s.t. } \int_0^1 (\theta - \mu - 2b)q(\theta)f(\theta)d\theta \geq 0. \end{aligned}$$

Solving this problem, $q(\theta) = 1$ if and only if

$$\theta - \mu + 2b + \xi(\theta - \mu - 2b) \geq 0,$$

where $\xi \geq 0$. The left hand side of the above equation is increasing in θ and equals to $4b \geq 0$ at $\theta = \mu + 2b$. It follows that R would optimally chooses a q such that for some k , $q(\theta) = 1$ for $\theta \geq k$ and $q(\theta) = 0$ for $\theta < k$. ■

Proposition 1

Proof of Proposition 1. First, to let C approve implementing the initiated policy when $s = 1$, R has to ensure that $\mathbb{E}[\theta|\theta \geq k] - b \geq \mu + b$ or, equivalently, $\mathbb{E}[\theta|\theta \geq k] \geq \mu + 2b$. Second, for any b and k such that $\mathbb{E}[\theta|\theta \geq k] \geq \mu + 2b$, R's expected payoff is

$$u_R(k, b, \mu) := \mu - b + \int_k^1 (\theta - (\mu - 2b)) f(\theta) d\theta,$$

which is strictly increasing in k if $k < \mu - 2b$ and strictly decreasing in k if $k > \mu - 2b$. Hence, if $\mathbb{E}[\theta|\theta \geq \mu - 2b] \geq \mu + 2b$, it is optimal for R to set $k = \mu - 2b$; and if $\mathbb{E}[\theta|\theta \geq \mu - 2b] < \mu + 2b$, it is optimal for R to set k as small as possible to satisfy $\mathbb{E}[\theta|\theta \geq k] \geq \mu + 2b$. Therefore, R's optimal experiment is characterized by

$$k^*(\mu, b) = \min \{k \geq \mu - 2b : \mathbb{E}[\theta|\theta \geq k] \geq \mu + 2b\}.$$

Note that $\mathbb{E}[\theta|\theta \geq \mu - 2b] - (\mu + 2b)$ is strictly decreasing in b . Moreover,

$$\begin{aligned} & \mathbb{E}[\theta|\theta \geq \mu] - \mu > 0 \\ \mathbb{E} \left[\theta|\theta \geq \mu - 2 \times \frac{1 - \mu}{2} \right] - (\mu + 2 \times \frac{1 - \mu}{2}) &= \mathbb{E}[\theta|\theta \geq 2\mu - 1] - 1 < 0. \end{aligned}$$

Hence, there exists a unique $\hat{b}(\mu) \in (0, \frac{1-\mu}{2})$ such that $\mathbb{E}[\theta|\theta \geq \mu - 2b] < \mu + 2b$ if and only if $b > \hat{b}(\mu)$. It follows that $k^*(b, \mu) = \mu - 2b$ if $b \leq \hat{b}(\mu)$. Furthermore, if $b > \hat{b}(\mu)$,

$$k^*(\mu, b) = \hat{k}(\mu, b) := \min \{k \geq 0 : \mathbb{E}[\theta|\theta \geq k] \geq \mu + 2b\},$$

which is increasing in b . ■

Proposition 2

The following two lemmas help to prove Proposition 2.

Lemma A.1. For $\hat{b}(\mu)$,

1. If $\mu \leq \frac{\mathbb{E}[\theta]}{2}$, $\hat{b}(\mu) = \frac{\mu}{2}$ and $\hat{k}(\mu + 2\hat{b}(\mu)) = 0$;
2. If $\mu > \frac{\mathbb{E}[\theta]}{2}$, $\hat{b}(\mu) < \frac{\mu}{2}$ and $\hat{k}(\mu + 2\hat{b}(\mu)) > 0$.

Proof. First, $\hat{k}(\mu + 2b) > \hat{k}(2\mu) \geq 0 > \mu - 2b$ for all $b > \frac{\mu}{2}$. Hence, $\hat{b}(\mu) \leq \frac{\mu}{2}$. Suppose $\mu \leq \frac{\mathbb{E}[\theta]}{2}$. Then, for any $b < \frac{\mu}{2}$, $\hat{k}(\mu + 2b) \leq \hat{k}(2\mu) = 0 < \mu - 2b$. Hence, $\hat{b}(\mu) \geq \frac{\mu}{2}$. It follows that $\hat{b}(\mu) = \frac{\mu}{2}$ and $\hat{k}(\mu + 2\hat{b}(\mu)) = \hat{k}(2\mu) = 0$. Now suppose $\mu > \frac{\mathbb{E}[\theta]}{2}$. Then, $\hat{k}(\mu + 2 \times \frac{\mu}{2}) = \hat{k}(2\mu) > 0 = \mu - 2 \times \frac{\mu}{2}$. Hence, $\hat{b}(\mu) < \frac{\mu}{2}$. It follows that $\hat{k}(\mu + 2\hat{b}(\mu)) = \mu - 2\hat{b}(\mu) > 0$. ■

Lemma A.2. $\hat{b}(\mu) < \frac{\mathbb{E}[\theta] - \mu}{2}$ if and only if $\mu < \frac{\mathbb{E}[\theta]}{2}$.

Proof. First, suppose $\mu < \frac{\mathbb{E}[\theta]}{2}$. Then, $\hat{b}(\mu) \leq \frac{\mu}{2} < \frac{\mathbb{E}[\theta] - \mu}{2}$. Second, suppose $\mu > \frac{\mathbb{E}[\theta]}{2}$. Then, $\hat{k}(\mu + 2\hat{b}(\mu)) > 0$, which implies that $\mu + 2\hat{b}(\mu) > \mathbb{E}[\theta]$, so that $\hat{b}(\mu) > \frac{\mathbb{E}[\theta] - \mu}{2}$. Third, suppose $\mu = \frac{\mathbb{E}[\theta]}{2}$. Then, $\hat{b}(\mu) = \frac{\mu}{2} = \frac{\mathbb{E}[\theta]}{4} = \frac{\mathbb{E}[\theta] - \mu}{2}$. ■

Proof of Proposition 2. By Lemma A.1 and Lemma A.2,

$$k^*(b, \mu) := \begin{cases} \mu - 2b, & b < \hat{b}(\mu) \\ 0, & \hat{b}_1(\mu) \leq b \leq \frac{\mathbb{E}[\theta] - \mu}{2} \\ \hat{k}(\mu + 2b), & b \geq \max\left\{\hat{b}(\mu), \frac{\mathbb{E}[\theta] - \mu}{2}\right\} \end{cases}$$

and $\hat{b}(\mu) < \frac{\mathbb{E}[\theta] - \mu}{2}$ if and only if $\mu < \frac{\mathbb{E}[\theta]}{2}$. Let

$$u_R^*(b, \mu) := \mu - b + \int_{k^*(b, \mu)}^1 (\theta - \mu + 2b)f(\theta)d\theta.$$

Then,

$$u_R^*(b, \mu) = \begin{cases} \mu - b + \int_{\mu - 2b}^1 (\theta - \mu + 2b)f(\theta)d\theta, & b < \hat{b}(\mu) \\ \mathbb{E}[\theta] + b, & \hat{b}(\mu) \leq b \leq \frac{\mathbb{E}[\theta] - \mu}{2} \\ \mu + b \left(3 - 4F\left(\hat{k}(\mu + 2b)\right)\right), & b \geq \max\left\{\hat{b}(\mu), \frac{\mathbb{E}[\theta] - \mu}{2}\right\}. \end{cases}$$

First, suppose $\mu < \frac{\mathbb{E}[\theta]}{2}$ and $\hat{b}(\mu) \leq b \leq \frac{\mathbb{E}[\theta] - \mu}{2}$. Then, $k^*(b, \mu) = 0$ and $u_R^*(b, \mu) = \mathbb{E}[\theta] + b > \mu$, so that R would initiate a policy and implement it through a “big bang.” This proves the second claim of the proposition.

Note that $u_R^*(b, \mu)$ is continuous in b and

$$u_R^*(0, \mu) = \mu + \int_{\mu}^1 (\theta - \mu) f(\theta) d\theta > \mu$$

$$u_R^*\left(\frac{1 - \mu}{2}, \mu\right) = \mu - \frac{1 - \mu}{2} < \mu.$$

Moreover, $u_R^*(b, \mu)$ is differentiable in b except for at most two points and

$$\frac{\partial}{\partial b} u_R^*(b, \mu) = \begin{cases} 1 - F(\mu - 2b), & b < \hat{b}(\mu) \\ 1, & \hat{b}(\mu) < b < \frac{\mathbb{E}[\theta] - \mu}{2} \\ 4 \left(1 - F\left(\hat{k}(\mu + 2b)\right)\right) \left(1 - \frac{2b}{\mu + 2b - \hat{k}(\mu + 2b)}\right) - 1, & b > \max\left\{\hat{b}(\mu), \frac{\mathbb{E}[\theta] - \mu}{2}\right\}. \end{cases}$$

Note that $1 - F(\mu - 2b)$ is strictly increasing in b and

$$4 \left(1 - F\left(\hat{k}(\mu + 2b)\right)\right) \left(1 - \frac{2b}{\mu + 2b - \hat{k}(\mu + 2b)}\right) - 1$$

is strictly decreasing in b . It follows that there are three possible cases with regard to the shape of $u_R^*(b, \mu)$: (1) $u_R^*(b, \mu)$ is strictly decreasing in b ; (2) $u_R^*(b, \mu)$ is inverse U-shaped in b and admits a maximum in $b \in \left[\hat{b}(\mu), \frac{1 - \mu}{2}\right]$; (3) $u_R^*(b, \mu)$ is U-shaped and admits a local minimum in $b \in \left[0, \hat{b}(\mu)\right]$ and inverse U-shaped in $b \in \left[\hat{b}(\mu), \frac{1 - \mu}{2}\right]$. Obviously, the first claim of proposition holds in the first two cases. In the third case, the minimum of $u_R^*(b, \mu)$ in $b \in \left[0, \hat{b}(\mu)\right]$ is attained at

$$b = x := \frac{\mu - F^{-1}\left(\frac{1}{2}\right)}{2}.$$

Hence, this case necessitates $x \in \left(0, \hat{b}(\mu)\right)$. It follows that for all $b \in \left[0, \hat{b}(\mu)\right]$,

$$u_R^*(b, \mu) \geq u_R^*(x, \mu) = \mu + \int_{\mu - 2x}^1 (\theta - \mu) f(\theta) d\theta > \mu + \int_{\mu - 2x}^1 (\theta - \mu - 2x) d\theta \geq \mu,$$

where the second equality is due to $F(\mu - 2x) = \frac{1}{2}$ and the last inequality is due to $x < \hat{b}(\mu)$. Because $u_R^*(b, \mu)$ is inverse U-shaped in $b \in \left[\hat{b}(\mu), \frac{1 - \mu}{2}\right]$, there exists a

unique $\bar{b}(\mu) \in \left(\hat{b}(\mu), \frac{1-\mu}{2}\right)$ such that $u_R^*(\bar{b}(\mu), \mu) = \mu$ and $u_R^*(b, \mu) \geq \mu$ if and only if $b \leq \bar{b}(\mu)$. ■

Proposition 3

Proof of Proposition 3. This proposition can be proved by several steps.

Step 1. The objective of this step is to show that if b is optimal, then either $b = 0$ or $b \in \left[\hat{b}(\mu), \frac{1-\mu}{2}\right]$. To see this, note that for any $b < \hat{b}(\mu)$, $k^*(b, \mu) = \mu - 2b$ and

$$u_R^*(b, \mu) = \mu - b + \int_{\mu-2b}^1 (\theta - \mu + 2b)f(\theta)d\theta.$$

It follows that

$$\frac{\partial^2}{\partial b^2} u_R^*(b, \mu) = 4f(\mu - 2b) > 0,$$

so that $u_R^*(b, \mu)$ is strictly convex in $b \in [0, \hat{b}(\mu)]$. This implies that

$$\arg \max_{b \in [0, \hat{b}(\mu)]} u_R^*(b, \mu) \in \{0, \hat{b}(\mu)\},$$

which proves the claim.

Let

$$U_1(\mu) := u_R^*(0, \mu) = \mu + \int_{\mu}^1 (\theta - \mu)f(\theta)d\theta$$

$$U_2(\mu) := \max_{b \in [\hat{b}(\mu), \frac{1-\mu}{2}]} u_R^*(b, \mu) = \mu - b + \int_{k^*(b, \mu)}^1 (\theta - \mu + 2b)f(\theta)d\theta.$$

On the basis of Step 1, the proposition can be proved by characterizing $U_2(\mu)$ and comparing it with $U_1(\mu)$.

Step 2. First, there exists a unique $\hat{\mu} \in \left(\frac{\mathbb{E}[\theta]}{2}, 1\right)$ that solves

$$1 - 2F\left(\mu - 2\hat{b}(\mu)\right) = 0$$

and $1 - 2F(\mu - 2\hat{b}(\mu)) > 0$ if and only if $\mu < \hat{\mu}$. Second, $U_2(\mu) < U_1(\mu)$ for all $\mu \geq \hat{\mu}$.

The first claim is due to the fact that $1 - 2F(\mu - 2\hat{b}(\mu))$ is strictly decreasing in $\mu \in \left[\frac{\mathbb{E}[\theta]}{2}, 1\right]$, equals to 1 at $\mu = \frac{\mathbb{E}[\theta]}{2}$, and equals to -1 at $\mu = 1$.

For the second claim, suppose $\mu \geq \hat{\mu} > \frac{\mathbb{E}[\theta]}{2}$. Then, $\hat{b}(\mu) > \frac{\mu}{2} > \frac{\mathbb{E}[\theta] - \mu}{2}$. It follows that for any $b < \hat{b}(\mu)$,

$$\frac{\partial}{\partial b} u_R^*(b, \mu) = 1 - 2F(\mu - 2b) < 1 - 2F(\mu - 2\hat{b}(\mu)) \leq 0;$$

and for any $b > \hat{b}(\mu)$, $k^*(b, \mu) = \hat{k}(\mu + 2b) > 0$ and

$$\begin{aligned} \frac{\partial}{\partial b} u_R^*(b, \mu) &= 1 - 2F(\hat{k}(\mu + 2b)) - 2f(\hat{k}(\mu + 2b)) (\hat{k}(\mu + 2b) - (\mu - 2b)) \hat{k}'(\mu + 2b) \\ &< 1 - 2F(\hat{k}(\mu + 2b)) \\ &< 1 - 2F(\mu - 2\hat{b}(\mu)) \leq 0. \end{aligned}$$

Then, by the continuity of $u_R^*(b, \mu)$ in b , $u_R^*(b, \mu)$ is strictly decreasing in $b \in [0, \frac{1-\mu}{2}]$. Hence, it must be true that $U_2(\mu) < U_1(\mu)$.

Step 3. The goal in this step is to prove $U_2(0) > U_1(0)$. Note that $U_1(0) = \mathbb{E}[\theta]$. But because $\hat{b}(0) = 0$,

$$\begin{aligned} U_2(\mu) &\geq u_R^*\left(\frac{\mathbb{E}[\theta]}{2}, \mu\right) \\ &= \mu - \frac{\mathbb{E}[\theta]}{2} + \int_{k^*(\mathbb{E}[\theta])}^1 (\theta + \mathbb{E}[\theta]) f(\theta) d\theta \\ &= \mu + \frac{3}{2}\mathbb{E}[\theta] > U_1(\mu), \end{aligned}$$

where the third inequality is due to $k^*(\mathbb{E}[\theta]) = 0$.

Step 4. Suppose $\mu \in (0, \hat{\mu})$. There exists a unique $\beta_R(\mu) \in (\hat{b}(\mu), b^*(\mu))$ that solves

$$\max_{b \in [\hat{b}(\mu), \frac{1-\mu}{2}]} u_R^*(b, \mu).$$

To show this, first note that for any b such that $\hat{b}(\mu) < b < \max\left\{\hat{b}(\mu), \frac{\mathbb{E}[\theta] - \mu}{2}\right\}$,

$\frac{\partial}{\partial b} u_R^*(b, \mu) = 1$; and for any $b > \max \left\{ \hat{b}(\mu), \frac{\mathbb{E}[\theta] - \mu}{2} \right\}$,

$$\begin{aligned} \frac{\partial}{\partial b} u_R^*(b, \mu) &= 4 \left(1 - F \left(\hat{k}(\mu + 2b) \right) \right) \left(\frac{\mu - \hat{k}(\mu + 2b)}{\mu + 2b - \hat{k}(\mu + 2b)} \right) - 1 \\ &= 4 \left(1 - F \left(\hat{k}(\mu + 2b) \right) \right) \left(1 - \frac{2b}{\mu + 2b - \hat{k}(\mu + 2b)} \right) - 1 \end{aligned}$$

is strictly decreasing in b . Then, by the continuity of $u_R^*(b, \mu)$ in b , $u_R^*(b, \mu)$ is strictly quasi-concave in $b \in \left[\hat{b}(\mu), \frac{1-\mu}{2} \right]$. Besides,

$$\frac{\partial}{\partial b} u_R^*(b^*(\mu), \mu) = -1$$

and

$$\lim_{b \downarrow \max \left\{ \hat{b}(\mu), \frac{\mathbb{E}[\theta] - \mu}{2} \right\}} \frac{\partial}{\partial b} u_R^*(b, \mu) = \begin{cases} 1, & \hat{b}(\mu) \leq \frac{\mathbb{E}[\theta] - \mu}{2} \\ 1 - 2F \left(\mu - 2\hat{b}(\mu) \right) > 0, & \hat{b}(\mu) > \frac{\mathbb{E}[\theta] - \mu}{2}. \end{cases}$$

Hence, there exists a unique $\beta_R(\mu) \in \left(\max \left\{ \hat{b}(\mu), \frac{\mathbb{E}[\theta] - \mu}{2} \right\}, b^*(\mu) \right)$ that solves

$$\frac{\partial}{\partial b} u_R^*(b, \mu) = 4 \left(1 - F \left(\hat{k}(\mu + 2b) \right) \right) \left(\frac{\mu - \hat{k}(\mu + 2b)}{\mu + 2b - \hat{k}(\mu + 2b)} \right) - 1 = 0$$

and maximizes $u_R^*(b, \mu)$.

At last, for any $\mu \in (0, \hat{\mu})$, $U_2(\mu) = u_R^*(\beta_R(\mu), \mu)$. Because $\beta_R(\mu) < b^*(\mu)$,

$$\hat{k}(\mu + 2\beta_R(\mu)) < \hat{k}(\mathbb{E}[\theta | \theta \geq \mu]) = \mu.$$

Hence, by the envelope theorem,

$$U_2'(\mu) - U_1'(\mu) = -f(\hat{k}) \left(\hat{k} - (\mu - 2\beta_R) \right) \hat{k}' - \left(F(\mu) - F(\hat{k}) \right) < 0,$$

where $\hat{k} = \hat{k}(\mu + \beta_R(\mu))$ and $\beta_R = \beta_R(\mu)$ for short. To summarize, $U_2(\mu) - U_1(\mu)$ is strictly decreasing in $\mu \in [0, \hat{\mu}]$, $U_2(0) - U_1(0) > 0$, and $U_2(\hat{\mu}) - U_1(\hat{\mu}) < 0$. Therefore, there exists a unique $\mu_R^* \in (0, \hat{\mu})$ such that $U_2(\mu_R^*) - U_1(\mu_R^*) = 0$ and $U_2(\mu) - U_1(\mu) > 0$ if and only if $\mu < \mu_R^*$.

Consequently, the optimal political stake is unique and equals to

$$b_R^*(\mu) := \begin{cases} \beta_R(\mu) \in (0, b^*(\mu)), & \mu < \mu_R^* \\ 0, & \mu \geq \mu_R^*. \end{cases}$$

This completes the proof. ■

Proposition 4

Proof of Proposition 4. By Proposition 1, Lemma A.1, and Lemma A.2,

$$u_C^*(b, \mu) = \begin{cases} \mu + b + \int_{\mu-2b}^1 (\theta - \mu - 2b) f(\theta) d\theta, & b < \hat{b}_1(\mu) \\ \mathbb{E}[\theta] - b, & \hat{b}(\mu) \leq b \leq \frac{\mathbb{E}[\theta] - \mu}{2} \\ \mu + b, & b \geq \max \left\{ \hat{b}(\mu), \frac{\mathbb{E}[\theta] - \mu}{2} \right\}. \end{cases}$$

This proposition can be proved through the following steps.

Step 1. If $\mu > F^{-1}(\frac{1}{2})$, then it is optimal for C to have some $b > 0$. To show this, note that for $b < \hat{b}_1(\mu)$,

$$\begin{aligned} \frac{\partial}{\partial b} u_C^*(b, \mu) &= 1 - 2(1 - F(\mu - 2b)) + 2(\mu - 2b - \mu - 2b) f(\mu - 2b) \\ &= 2F(\mu - 2b) - 1 - 8bf(\mu - 2b). \end{aligned}$$

Hence, if $\mu > F^{-1}(\frac{1}{2})$,

$$\frac{\partial}{\partial b} u_C^*(0, \mu) = 2F(\mu) - 1 > 0,$$

so that $u_C^*(\epsilon, \mu) > u_C^*(0, \mu)$ for a sufficiently small $\epsilon > 0$.

Step 2. If $\mu \leq F^{-1}(\frac{1}{2})$, then $\bar{b}(\mu) > \max \left\{ \hat{b}(\mu), \frac{\mathbb{E}[\theta] - \mu}{2} \right\}$. To show this, first, assume $\mu < \frac{\mathbb{E}[\theta]}{2}$, so that $\max \left\{ \hat{b}(\mu), \frac{\mathbb{E}[\theta] - \mu}{2} \right\} = \frac{\mathbb{E}[\theta] - \mu}{2}$ and $k^* \left(\frac{\mathbb{E}[\theta] - \mu}{2}, \mu \right) = 0$. Then,

$$u_R^* \left(\frac{\mathbb{E}[\theta] - \mu}{2}, \mu \right) = \mathbb{E}[\theta] + \frac{\mathbb{E}[\theta] - \mu}{2} > \mu.$$

Hence, by the definition of $\bar{b}(\mu)$, $\max \left\{ \hat{b}(\mu), \frac{\mathbb{E}[\theta] - \mu}{2} \right\} < \bar{b}(\mu)$. Second, assume $\mu \geq \frac{\mathbb{E}[\theta]}{2}$,

so that $\max \left\{ \hat{b}(\mu), \frac{\mathbb{E}[\theta] - \mu}{2} \right\} = \hat{b}(\mu)$ and

$$k^* \left(\hat{b}(\mu), \mu \right) = \hat{k} \left(\mu + 2\hat{b}(\mu) \right) = \mu - 2\hat{b}(\mu).$$

Then,

$$u_R^* \left(\hat{b}(\mu), \mu \right) = \mu + \hat{b}(\mu) \left(3 - 4F \left(\mu - 2\hat{b}(\mu) \right) \right).$$

But because $\mu \leq F^{-1} \left(\frac{1}{2} \right)$, $\mu - 2\hat{b}(\mu) < F^{-1} \left(\frac{3}{4} \right)$. It follows that

$$3 - 4F \left(\mu - 2\hat{b}(\mu) \right) > 0$$

and so $u_R^* \left(\hat{b}(\mu), \mu \right) > \mu$. Hence, by the definition of $\bar{b}(\mu)$, $\max \left\{ \hat{b}(\mu), \frac{\mathbb{E}[\theta] - \mu}{2} \right\} < \bar{b}(\mu)$.

Step 3. There exists a unique $\mu_C^* \in \left[0, F^{-1} \left(\frac{1}{2} \right) \right]$ such that for $\mu \leq F^{-1} \left(\frac{1}{2} \right)$,

$$\int_{\mu}^1 (\theta - \mu) f(\theta) d\theta \geq \bar{b}(\mu)$$

if and only if $\mu \leq \mu_C^*$. Moreover, $\mu_C^* > 0$ if and only if

$$\int_{F^{-1} \left(\frac{3}{4} \right)}^1 \theta f(\theta) d\theta < \int_0^{F^{-1} \left(\frac{3}{4} \right)} \theta f(\theta) d\theta.$$

To prove this, let

$$x(\mu) := \int_{\mu}^1 (\theta - \mu) f(\theta) d\theta.$$

First, for $\mu < F^{-1} \left(\frac{1}{2} \right)$,

$$\frac{d}{d\mu} \left(x(\mu) - \frac{1 - \mu}{2} \right) = F(\mu) - \frac{1}{2} < 0$$

and so $x(\mu) - \frac{1-\mu}{2}$ is strictly decreasing in $\mu < F^{-1}\left(\frac{1}{2}\right)$. Second,

$$\begin{aligned} x\left(F^{-1}\left(\frac{1}{2}\right)\right) &= \int_{F^{-1}\left(\frac{1}{2}\right)}^1 \theta f(\theta) d\theta - \frac{F^{-1}\left(\frac{1}{2}\right)}{2} \\ &< 1 - F\left(F^{-1}\left(\frac{1}{2}\right)\right) - \frac{F^{-1}\left(\frac{1}{2}\right)}{2} \\ &= \frac{1 - F^{-1}\left(\frac{1}{2}\right)}{2}. \end{aligned}$$

Hence, there exists a unique $\mu_0 \in [0, F^{-1}\left(\frac{1}{2}\right))$ such that $x(\mu) \leq \frac{1-\mu}{2}$ if and only if $\mu \geq \mu_0$. Note that because $x(0) = \mathbb{E}[\theta]$, $\mu_0 = 0$ if and only if $\mathbb{E}[\theta] \geq \frac{1}{2}$.

Consider $\mu \geq \mu_0$, so that $x(\mu) \leq \frac{1-\mu}{2}$. By the definition of $\bar{b}(\mu)$, $x(\mu) > \bar{b}(\mu)$ if and only if

$$u_R^*(x(\mu), \mu) - \mu = x(\mu) \left(3 - 4F\left(\hat{k}(\mu + 2x(\mu))\right)\right) < 0,$$

or, equivalently, if and only if

$$\hat{k}(\mu + 2x(\mu)) > F^{-1}\left(\frac{3}{4}\right).$$

Note that for $\mu < F^{-1}\left(\frac{1}{2}\right)$,

$$\frac{d}{d\mu}(\mu + 2x(\mu)) = 2F(\mu) - 1 < 0$$

and so $\hat{k}(\mu + 2x(\mu))$ is strictly decreasing in $\mu < F^{-1}\left(\frac{1}{2}\right)$.

First, $\hat{k}(2x(0)) = \hat{k}(2\mathbb{E}[\theta]) > F^{-1}\left(\frac{3}{4}\right)$ if and only if

$$\int_{F^{-1}\left(\frac{3}{4}\right)}^1 (\theta - 2\mathbb{E}[\theta]) f(\theta) d\theta = \int_{F^{-1}\left(\frac{3}{4}\right)}^1 \theta f(\theta) d\theta - \frac{1}{2} \int_0^1 \theta f(\theta) d\theta < 0,$$

or, equivalently, if and only if

$$\int_{F^{-1}\left(\frac{3}{4}\right)}^1 \theta f(\theta) d\theta < \int_0^{F^{-1}\left(\frac{3}{4}\right)} \theta f(\theta) d\theta.$$

If this condition fails, $\hat{k}(\mu + 2x(\mu)) < F^{-1}(\frac{3}{4})$ and so $x(\mu) < \bar{b}(\mu)$ for all $\mu \leq F^{-1}(\frac{1}{2})$. Suppose this condition holds. Then, there exists a unique $\mu_C^* \in (0, F^{-1}(\frac{1}{2})]$ such that $\hat{k}(\mu + 2x(\mu)) \geq F^{-1}(\frac{3}{4})$ and so $x(\mu) \geq \bar{b}(\mu)$ if and only if $\mu \leq \mu_C^*$. Note that $\mu_C^* = F^{-1}(\frac{1}{2})$ if and only if $\hat{k}(F^{-1}(\frac{1}{2}) + 2x(F^{-1}(\frac{1}{2}))) \geq F^{-1}(\frac{3}{4})$. This completes the proof of this step and therefore the proposition. ■