Unification versus Separation of Regulatory Institutions*

Dana Foarta
Stanford GSB

Takuo Sugaya
Stanford GSB

February 17, 2017

Abstract

Why might a country choose to aggregate regulatory information into a single government agency? And what might reverse this choice? We consider an oversight setting in which the institutional structure affects access to information. A regulator exerts effort towards a final outcome, but an oversight authority can intervene, which prevents the final outcome from being reached. We examine the choice between institutional unification and separation. Unification affords the oversight authority more information on the probable outcome. Yet dynamically, institutional separation improves learning about the quality of regulatory agencies. This leads to switching between unification and separation in equilibrium.

JEL Classification Codes: P16, D82, H11, G18.

*Foarta: Stanford Graduate School of Business, 655 Knight Way, Stanford, CA, 94305, tel: (650) 723-4058, email: ofoarta@stanford.edu. Sugaya: Stanford Graduate School of Business, 655 Knight Way, Stanford, CA, 94305, tel: (650) 724-3739, email: tsugaya@stanford.edu. We thank Steve Callander, Peter Buisseret, Ken Shotts, Andy Skrzypacz, Victoria Vanasco, as well as seminar partipants at Arizona State, Boston University, Caltech, LSE, Stanford, University of Arizona, University of Warwick, and Yale for useful comments and suggestions.

1 Introduction

A central question in political economy has been whether to unify government functions or to separate them into discrete institutions.[1] Of the various aspects of institutions that have come under focus, regulators' access to information stands out as a key element in policy debates. Information is the lifeblood of effective regulatory action. Not infrequently after crises comes the conclusion that information was overly dispersed, not aggregated, and that the crisis could have easily been avoided if only the officials had known. This was the case after the global financial crisis, and it was also the case after the terrorist attacks of 2001. It stands to reason, therefore, that the solution is to aggregate information into a single government agency. Indeed, this is exactly the response we witnessed after the great financial crisis of 2008 and the terrorist attacks of 2001, as well as after many smaller yet non-trivial regulatory failures.

If this were all that has happened we could conclude that these were episodes of flawed governments correcting organizational mistakes. Yet, the choice of separated regulatory structures–where information and authority is dispersed across multiple independent agencies–is widespread, so much so that it cannot be viewed as a mistake. Consider, for instance, financial regulation, specifically the functions of bank supervision and the lender of last resort. Bank supervisors monitor risk taking in the banking sector. Their effort in monitoring is meant to reduce excessive risk-taking by banks, which in turn reduces the probability of banking crises. The lender of last resort relies on the work of the bank supervisors. If there are signs of a banking crisis, it can intervene to prevent a crisis from unfolding. Therefore, these two functions are closely linked by their role in preventing banking crises. Figure 1 illustrates the considerable heterogeneity across countries in the choice of whether to unify or separate these two functions.[2]

[1] See Hayek (1945), Tiebout (1956), Oates (1972), Weingast (1995), Blanchard and Shleifer (2001), Besley and Coate (2003a), Cai and Treisman (2005), Bardhan and Mookherjee (2006) and the works referenced within.

[2] Data from "How countries supervise their banking, insurers and securities markets," Central Banking Publications for year 2013 and from Melecky and Podpiera (2013) for years 1999-2010. It covers 98 countries. The data for years 2011-2012 was not available.

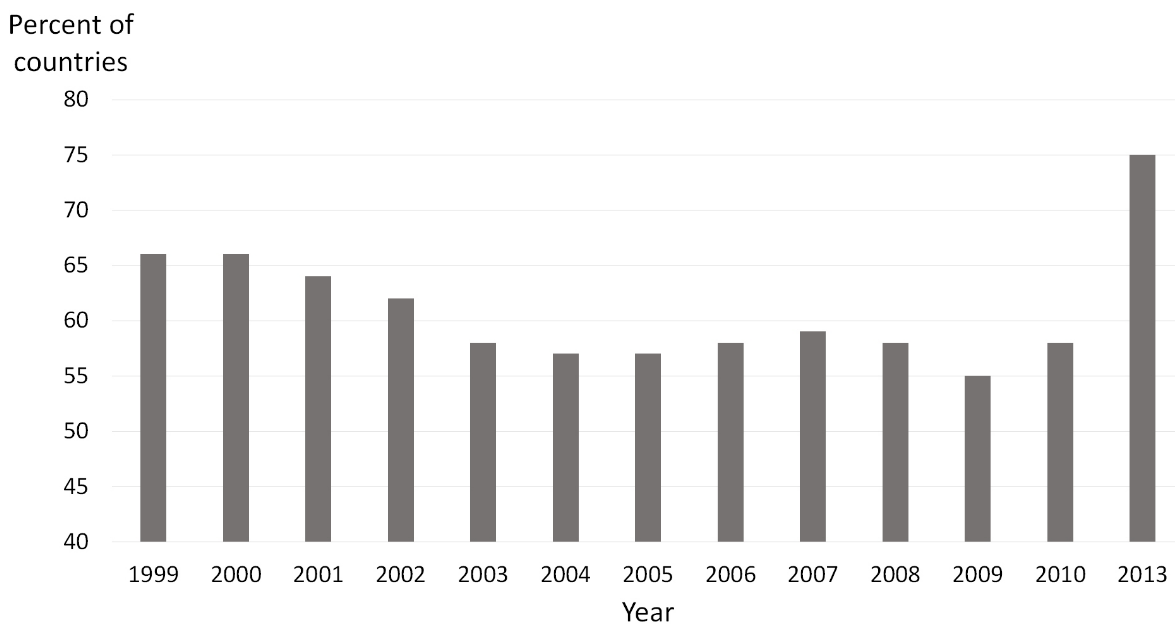

Figure 1: Percent of countries with institutional unification of bank supervision inside the central bank

Figure 1 also reveals a second puzzle. The choices of regulatory structure vary over time, and not in a way that is purely reactive to crises. Indeed, changes go both ways, from a separated to a unified structure and the other way around. In fact, a single country can exhibit both types of change. The United Kingdom, for instance, moved from a system with institutional unification to one with separation in 1997/98. It created the Financial Services Authority (FSA) as an institution separated from the Bank of England, which retained the lender-of-last-resort function. Yet, in 2013, the UK switch back again, with the responsibility for bank supervision transferred back and unified within the Bank of England.

Why might a country choose a separated rather than a unified regulatory structure when information flows are affected? And what might reverse this choice? In this paper, we develop a model of regulatory oversight to answer these two questions and to make sense of these empirical patterns. We take bank regulation as our leading example, although our model is applicable to any oversight setting in which the flow of information is important.

In our model, a regulatory agency (referred to as 'the regulator') exerts costly effort to increase the probability of a good outcome. How costly the effort is depends on the regulator's private type. After the regulator supplies effort and before the final outcome is revealed, an overseer–the central bank in our motivating example–can make a costly intervention–it can intervene as a lender of last resort. This action stops the final outcome from being reached–everyone observes the intervention, and no one knows what the final outcome would have been. The cost of the intervention is lower than the cost of the bad outcome, but the good outcome carries no cost at all. Therefore, an overseer who aims to minimize costs wants to intervene only if the final outcome is likely to be bad. In a unified institutional structure, the overseer has access to additional information about the likely outcome: she observes a noisy signal about the final outcome. In a separated institutional structure, such a signal is not available. In this environment, we consider the problem of selecting quality regulators given the problems of moral hazard and adverse selection–the citizens only observe the final outcome, and based on it they form a belief about the regulator's quality. We implicitly assume that the citizens' goal is shared by the government, who conducts the regulatory reforms in order to achieve this goal. In a dynamic setting, this sequence of actions is repeated every period, and the government can each time choose to replace the regulator by overhauling the regulatory agency. Therefore, the regulatory agency has reputational concerns, as its continuation depends on the beliefs formed about it.

We first identify why, in a static sense, unified oversight may not be optimal. Unification allows the overseer access to additional information about the final outcome. Yet, the overseer may use this information to intervene not only when it is necessary, but also when no intervention is needed. In our motivating example, a lender of last resort may bail out banks, but it may also bail out the supervisors for their lack of effort in oversight. This potential for over-intervention reduces the incentive for the regulator to exert effort in the first place. The opposite happens with institutional separation–the lack of information removes the possibility of unnecessary interventions and increases the incentive for the regulator to exert effort.

This result suggests why countries may make different choices but it does not explain the variance over time–why, for instance, the UK switched back and forth their regulatory structure. We expand our static analysis to a fully dynamic model and we show that dynamically, the informational advantage of unification comes at a higher price, because it limits the government's ability to learn about regulator quality. Precisely because unification leads to over-intervention, citizens (and the government) more often do not observe what outcome would have been reached had there not been an intervention. This worsens the government's ability to screen regulators, since they are not able to learn as much about the quality of the regulator from observing an intervention rather than the final outcome. Less effective screening leads, in turn, to low quality regulatory agencies being kept unreformed for longer. Thus, the static informational advantage of unification comes with a dynamic inefficiency. A trade-off emerges between using information to immediately correct potential regulatory failures–through the overseer's intervention–and waiting more often until the final outcome in order to learn about the regulator's quality.

We fully characterize the best sustainable equilibrium for citizens, and we show that the interplay between the static and the dynamic trade-offs leads to a changing regulatory structures in the optimal design of institutions. The dynamic benefit of better screening is highest when the regulator's reputation is intermediate–when there is enough uncertainty about the regulator's quality. When the regulator's reputation is very low or very high, observing the final outcome brings little benefit. In fact, as citizens (and the government) become convinced that the regulator is of high quality, the benefit of screening vanishes. This dynamic leads to switches in institutional structure between unification when the regulator's reputation is low, separation when the regulator's reputation is intermediate, and then unification again when the regulator's reputation becomes sufficiently high.

The results of the model suggest that transitions between institutional separation and unification can be welfare improving in environments in which society must balance incentives for regulatory effort and effective oversight. Our explanation resonates with the data presented in Figure 1 and with anecdotal accounts of these institutional changes. For example, when the UK regulatory agency–the FSA–was separated in 1997/1998, the main argument provided for this change was that it would "reduc[e] the risk that the central bank would be too accommodative in order to help frail intermediaries."[3] Similarly, the 2009 review of the FSA mentioned that, while unification would bring the benefit of "ensuring that banking liquidity supervision and central bank liquidity operations are closely linked, there is also a danger that this can reduce supervisory discipline."[4] When the supervisory functions were eventually transferred back to the Bank of England, the main reason cited was the lack of shared information between FSA and the Bank of England during the financial crisis.[5] Finally, going beyond this particular example, perceptions of high regulator quality have been shown to be important for market outcomes and for regulatory reform decisions.[6]

Related Literature. The questions our paper addresses are related to the broader debate in political economy on whether to unify government functions or to separate them into discrete institutions. By addressing unification versus separation in the context of regulatory oversight, we contribute to the literature on optimal institutional structure and the optimal degree of transparency in public institutions (Aghion and Tirole, 1997; Persson et al., 1997; Besley and Coate, 2003b; Maskin and Tirole, 2004; Prat, 2005; Alesina and Tabellini, 2007, among many others). Our paper focuses on a specific type of institutional relationship, that in which one institution has oversight power over another, and the overseer can take an action that reduces the informativeness of the observed outcome–after an intervention, no one knows what the outcome would have been otherwise. The institutional structure affects the information that the overseer has when making the intervention decision.

[3] The quote is from the executive summary of Chapter 14 of "Money, Banking, and Financial Markets" by Stephen G. Cecchetti and Kermit L. Schoenholtz, available at http://www.mhhe.com/economics/cecchetti/Cecchetti2_Ch14_RegDivofLabor.pdf

[4] FSA (2009), "The Turner Review," pp 92, http://www.fsa.gov.uk/pubs/other/turner_review.pdf , accessed December 5, 2016.

[5] For example, see "Responsibility without information: the Bank of England as Lender of Last Resort" in Financial Times, available at http://blogs.ft.com/maverecon/2007/09/responsibility-html/#axzz4M3r4Tqv9, accessed December 5, 2016.

[6] Restoring market confidence in regulators is often quoted as the reason for regulatory overhauls (see for example the above quoted Turner Review or the HM Treasury 2011 report, "A new approach to financial regulation"). More broadly, the importance of public perceptions and reputation has been shown to matter across regulatory areas (see for example Carpenter 2001, 2010).

Our paper is also related to the literature on the efficiency of oversight (De Mesquita and Stephenson, 2007; Fox and Van Weelden, 2010; Buisseret, 2016). While much of this literature has focused on static trade-offs in richer static models, we consider the dynamics of the regulatory structure. Our static model is simplified in order to focus on the dynamic effects of institutional choice–only the regulator has reputational concerns and the regulator works on only one task. We also focus on a different static trade-off than previous works in this literature–regulatory effort is discouraged when the overseer has access to information. The only other work we are aware of which examines this particular type of oversight relationship is Levitt and Snyder (1997). They address this tension in a static model with moral hazard. We consider the dynamic problem, with moral hazard and adverse selection, both of which are essential for our result of switches between unification and separation in the optimal institutional design.

Our paper also contributes to the literature on dynamic retention of public officials. Our model is related to Banks and Sundaram (1993, 1998) in that it incorporates both adverse selection and moral hazard in a dynamic setting with the possibility of removal of agents at the end of each period; however, we focus on the case in which both the regulator and the citizens are long-lived agents and in which the action of the overseer can reduce public information, which changes the dynamics of the problem. The result of switching between institutional unification and separation also relates this work to the literature on policy cycles (Ales et al., 2014; Dovis et al., 2016). As in this literature, we focus on the best sustainable equilibrium for the citizens. Our result of a changing institutional structure relies nevertheless on a distinct mechanism, coming from the interaction between the public's ability to learn about the agent's type and the agent's incentives to supply effort.

The rest of the paper is organized as follows. Section 2 describes the model. Section 3 presents static trade-off generated by the model in a one-period game. Section 4 presents the full dynamic model. Section 5 concludes, and the Appendix contains the proofs and extensions.

2 Environment

We begin by describing an environment that captures the main forces of the model. There are three players: a regulator (R),[7] an overseer (O), and a unitary public (P).

The Regulator. The regulator R has a type $\theta \in \{L, H\}$, which is private information. The type θ takes value H with commonly known probability μ. The regulator can exert unobservable effort $e \in [0, 1]$ to influence the probability distribution of a final outcome $y \in Y = \{B, G\}$. The probability of outcome $y = G$, denoted $q(e)$, is an increasing function of the effort e. The function $q(e)$ is twice continuously differentiable and concave:

$$q'(e) \geq 0,\ q''(e) \leq 0.$$

We also assume $q'(0) > 0$ since otherwise $q'(e) \geq 0$ and $q''(e) \leq 0$ would imply $q(e)$ is constant.

Exerting effort is costly, and the cost of implementing e for R of type θ is given by $c_\theta(e)$. The marginal cost of providing effort e is higher for type L than for type H. We assume that the functions $c_H(e)$ and $c_L(e)$ are twice continuously differentiable, $c'_\theta(0) = 0$, $\lim_{e \to 1} c'_\theta(e) = \infty\ \forall \theta$, and

$$c'_\theta(e) > 0,\ c''_\theta(e) > 0,\ c'_H(e) < c'_L(e)\ \ \forall e > 0. \tag{1}$$

The regulator receives a payoff that depends on his cost of effort $c_\theta(e)$ and on the expectation formed by the public about his type:

$$u_\theta = u^R(b) - c_\theta(e), \tag{2}$$

where $b \equiv (\mathbb{E}[\Pr(\theta = H | \textit{public outcome})])$, such that $u^R(b)$ is a function of the expectation formed by the public about R's type θ.[8]

[7] We consider a unitary regulatory agency, to which we refer to as 'the regulator'.

[8] The regulator's utility could also be modified to directly depend on the the public outcome, without

The utility function captures the benefit to the regulator from having a good reputation, as a bad reputation can translate into the regulatory agency being restructured in costly ways for the regulator.

The Overseer. After R exerts effort e, and before the final outcome y is revealed, the overseer O decides whether to intervene, $\iota \in \{0,1\}$, where $\iota = 1$ corresponds to an intervention and $\iota = 0$ corresponds to no intervention. Before the intervention decision is made, the institutional structure determines how much information the overseer has access to. In the separated institutional structure, the overseer has access to no information; in the unified institutional structure, the overseer receives a noisy signal $s \in S = \{B, G\}$ about the outcome y. For each effort e, the error in the signal is:

$$\Pr(s = B|y = G, e) = \Pr(s = G|y = B, e) = \varepsilon > 0.^{9} \tag{3}$$

We assume the following costs associated with the intervention decision ι and final outcome y:

$$C = \begin{cases} C_1 & \text{if} \quad \iota = 1, \\ 0 & \text{if} \quad \iota = 0 \text{ and } y = G, \\ C_2 & \text{if} \quad \iota = 0 \text{ and } y = B, \end{cases} \tag{4}$$

where $C_1, C_2 \in \mathbb{R}$, $0 < C_1 < C_2$. This specification captures the many situations in which early, preventive intervention is less costly than letting the situation worsen to a crisis; however, the actual cost savings happen only if the situation were to indeed worsen without intervention, otherwise costs would be saved by not performing an intervention. In our example of bank regulation, C_1 denotes the cost of an intervention to prevent a banking crisis when there are signs of a potential crisis, while $C_2 > C_1$ denotes the cost to society when a crisis is unfolding.

We make the following assumption about the costs C_1 and C_2:

changing the results of the model.

[9] All the results go through if the probability of error depends on y: $\Pr(s = B|y = G, e) = \varepsilon_G$ and $\Pr(s = G|y = B, e) = \varepsilon_B$.

Assumption 1 *The values C_1 and C_2 satisfy*

$$1+\frac{q\left(1\right)}{1-q\left(1\right)}\frac{\varepsilon}{1-\varepsilon}<\frac{C_2}{C_1}<1+\frac{q\left(0\right)}{1-q\left(0\right)}\frac{1-\varepsilon}{\varepsilon} \tag{5}$$

and

$$\frac{1}{1-q\left(0\right)}>\frac{C_2}{C_1}. \tag{6}$$

Expression (5) implies that it is less costly to choose $\iota=1$ after a signal $s=B$, regardless of R's effort. In addition, the expected cost of choosing $\iota=0$ is lower than the cost of $\iota=1$ whenever the signal is $s=G$, regardless of R's effort. Expression (6) implies that, if O does not obtain any information prior to her decision, then the expected cost of not intervening is lower than the cost of intervention.

Finally, O's objective is to minimize costs:

$$u^O=-C. \tag{7}$$

The Public. The public (P) is a unitary agent. It does not observe signal s. It only observes the cost paid by O according to (7). It can therefore infer the realization of y if $\iota=0$, but can only infer that $\iota=1$ if the cost is C_1. Based on these observations, the public updates its beliefs about R's type, following Bayes' rule whenever possible. In order to isolate the agency problem stemming from the regulator's reputational concern, we consider the case in which the public's utility is identical to that of the overseer:

$$u^P=-C. \tag{8}$$

In the interpretation of R and O as institutions in charge of banking stability, the public only observes a banking crisis–the bad outcome–if the lender of last resort did not intervene to prevent it. If the lender of last resort intervened, then the public does not know what would have happened without intervention–just isolated bank failures ($s=B$ but $y=G$) or a banking crisis ($y=B$).

We implicitly assume that the public's objective and information are shared by the government, so that the government is acting according to the public's preferences. The public's beliefs about the regulator then matter for the regulator because the government can impose costs on the regulatory agency though overhauls.

Timing. The model unfolds as follows:

1. Nature decides R's privately observed type $\theta \in \{L, H\}$, where $\Pr(\theta = H) = \mu$.

2. R supplies effort $e \in [0, 1]$ by paying cost $c_\theta(e)$. The effort e is R's private information.

3. Given e, the final outcome $y \in \{G, B\}$ and signal $s \in \{G, B\}$ are drawn from $\Pr(y, s|e)$.

4. In the unified institutional structure, O observes s. In the separated institutional structure, s is not produced; O does not observe y.

5. O decides $\iota \in \{0, 1\}$.

6. Cost C is observed by P, beliefs are updated, and payoffs are realized.

3 The Static Trade-off

We begin by analyzing the static trade-off induced by the institutional structure, and then present the model in a dynamic setting. Consider the following payoff function $u^R(b)$:

$$u^R(b) = \begin{cases} 1 & \text{if } \Pr(\theta = H|C) \geq \mu, \\ 0 & \text{if } \Pr(\theta = H|C) < \mu. \end{cases} \tag{9}$$

The payoff function $u^R(b)$ reflects the case in which R is removed whenever his reputation decreases–whenever the public's belief that R is a type H is lower than its prior. As it will be shown in the results of the next section, this payoff function comes out of optimal strategy for the public in the dynamic game.

We first analyze the static problem in each of the two institutional setups, separated and unified. Afterwards, we compare these two setups in terms of the effort provision and the expected costs to the public.

3.1 Separate Institutions

With separate institutions, no additional information becomes available for O before the intervention decision.

Definition 1 *An equilibrium with separate institutions is a set of effort levels e_θ and intervention policies ι such that (i) O chooses ι to maximize (7), (ii) e_θ maximizes (2), and (iii) $\Pr(\theta = H|C)$ is derived by P following Bayes' rule whenever possible.*

In the case with separate institutions, two equilibria exist: an uninteresting equilibrium in which $e_H = e_L = 0$,[10] and a more interesting equilibrium in which type H takes positive effort, and $e_H > e_L \geq 0$. In what follows, we analyze the latter equilibrium.

Proposition 1 *There exist a unique equilibrium with positive effort and separate institutions. The equilibrium has the following properties:*

- ***(Effort Strategy)*** *The equilibrium effort choices satisfy $e_H > e_L \geq 0$.*
- ***(Intervention Strategy)*** *O implements $\iota = 0$.*
- ***(Beliefs)*** *The equilibrium beliefs satisfy:*

$$\Pr(\theta = H|C = 0) > \mu > \Pr(\theta = H|C = C_2).$$[11]

Proof. In Appendix A.1.1. ■

[10] In the equilibrium with $e_H = e_L = 0$, the public belief is constant for each C. Hence, it is in turn optimal for the regulator to choose zero effort.

[11] We omit $\Pr(\theta = H|C = C_1)$ since $C = C_1$ never happens on the equilibrium path given the intervention strategy.

Intuitively, since type H can always pretend to be the type L with a lower cost, the former should obtain the higher value in equilibrium. If type H chose lower effort and enjoyed a higher expected reputation, then type L would deviate to type H's effort level in order to save on the effort cost.

Given (6), we conclude that O never intervenes since O does not obtain any additional information before intervention. In our motivating example, the central bank does not want to intervene, i.e., choose $\iota = 1$, unless it receives bad news, so the default policy is no intervention, $\iota = 0$.

Formally, regulatory effort e_θ is chosen by R in order to maximize his expected utility:

$$q(e_\theta)\, u^R(b|C = 0) + (1 - q(e_\theta))\, u^R(b|C = C_2) - c_\theta(e_\theta).$$ [12]

It follows then that reputation increases after a good final outcome, and it decreases after a bad final outcome. Given (9) and the first-order condition for effort, any positive effort choice satisfies:

$$q'(e_\theta) = c'_\theta(e_\theta). \tag{10}$$

3.2 One Unified Institution

We now consider the unified institutional structure. In this setting, the overseer O receives signal s about the final outcome before making the intervention decision.

Definition 2 *An equilibrium with a unified institutional structure is a set of effort levels e_θ and intervention policies $\iota(s)$ such that (i) O chooses $\iota(s)$ to maximize (7) conditional on each realization of s, (ii) e_θ maximizes (2), and (iii) $\Pr(\theta = H|C)$ is derived by P following Bayes' rule whenever possible.*

As in the case of separated institutions, there is always an equilibrium with no effort. For the rest of the discussion, we focus on the more interesting equilibrium in which at least type H exerts positive effort.

[12] Where $b|C$ means the public belief about θ given C.

Proposition 2 *There exits a unique equilibrium with positive effort and a unified institution. The equilibrium satisfies the following properties:*

- ***(Effort Strategy)*** *The equilibrium effort choices satisfy* $e_H > e_L \geq 0$.
- ***(Intervention Strategy)*** O *implements* $\iota = 1$ *after* $s = B$ *and* $\iota = 0$ *after* $s = G$.
- ***(Beliefs)*** *The equilibrium beliefs satisfy:*

$$\Pr(\theta = H|C = 0) > \mu > \Pr(\theta = H|C = C_1) \geq \Pr(\theta = H|C = C_2).$$

Proof. In Appendix A.1.2. ■

Proposition 2 shows that, in the equilibrium with $e_H > 0$, type H chooses strictly higher effort than type L. The intuition is the same as in the case with separated institutions: type H faces a lower marginal cost of effort, so he faces a higher relative gain from effort. The overseer O now observes the signal directly, and she uses that information to choose the cost-minimizing policy ι.

Compared to the case of separated institutions, the availability of information to O leads to a distortion to R's effort choice:

Corollary 1 (Effort withholding) *With institutional unification, R's effort choice is lower than in the case with institutional separation.*

Consider R's problem in the unified institutional structure. Given the public's equilibrium beliefs, the regulator's objective becomes

$$\max_{e_\theta} q\left(e_\theta\right)\left(1-\varepsilon\right) - c_\theta\left(e_\theta\right).$$

The first-order condition for positive effort is then

$$\left(1-\varepsilon\right) q'\left(e_\theta\right) = c'_\theta\left(e_\theta\right). \tag{11}$$

Notice that the effort corresponding to the separated institutional structure can be obtained in (11) if $\varepsilon = 0$. Let e_θ be the solution to (11). Applying the implicit function theorem to the first-order condition, we obtain

$$\frac{de_\theta}{d\varepsilon} = -\frac{q'(e_\theta)}{c''_\theta(e_\theta) - (1-\varepsilon) q''(e_\theta)} < 0, \tag{12}$$

which implies that e_θ decreases in the unified case compared to the separated case. The unified institutional structure reduces the ex-ante incentive to supply effort. Since O intervenes after $s = B$, this decreases the probability that the good outcome will be reached, reducing the expected reputational payoff to R from supplying effort.

3.3 Discussion of the Static Trade-off

Corollary 1 illustrates the main trade-off embedded in the choice of institutional structure: with institutional separation, R has the ex-ante incentive to put in more effort, but O's intervention decision is made without the benefit of additional information about the outcome; in the unified structure, the intervention decision is made with additional information, but the regulatory effort is reduced.

Let e^s_θ denote R's effort choice with separated institutions, and let e^ε_θ denote R's effort choice in the unified institution. The expected increase in costs when moving from the separated structure to the unified structure is given by

$$\begin{aligned} \Delta W(\varepsilon, C_2, C_1) \equiv & -\sum_\theta \Pr(\theta)\,(1 - q(e^s_\theta))\,C_2 \\ & + \sum_\theta \{\Pr(\theta)\,((1-\varepsilon)\,(1 - q(e^\varepsilon_\theta)) + \varepsilon q(e^\varepsilon_\theta))\,C_1 \\ & + \Pr(\theta)\varepsilon\,(1 - q(e^\varepsilon_\theta))\,C_2\}\,. \end{aligned} \tag{13}$$

The first term of the expression subtracts the expected cost under institutional separation. The second term adds the expected cost under institutional unification when the signal is $s = B$. Finally, the third term adds the expected cost under institutional unification when the signal is $s = G$ and the outcome is $y = B$.

The analysis of ΔW allows us to derive two key comparative statics. First, consider an increase in the relative cost C_1/C_2. That is, an intervention is less effective at reducing costs relative to the cost of a crisis. Then, the benefit of additional information brought about by unification becomes smaller. Second, consider an increase in the signal noise ε. This rends the signal less accurate, so the unified institution generates cost C_1 more often when the outcome would have been good. This type of unnecessary intervention diminishes the return to effort, because the regulator's reputation is hurt by intervention, as seen in (11). This in turn reduces R's effort provision.

Proposition 3 *The expected cost of the intervention policy ι is larger with institutional unification compared to institutional separation if ε and $\frac{C_1}{C_2}$ are sufficiently high.*

Proof. In Appendix A.1.3. ■

The above result highlights the fundamental reason why unification does not always dominate separation: the inability of O to commit not to choose $\iota = 1$ after $s = B$. With commitment, unification would always achieve an outcome at least as good as separation.

In Appendix A.2, we present two extensions of the static model. First, we show that the same trade-off emerges under a more general form of the regulator's utility function. Second, we show that the trade-off still exists even if in the unified institution the regulator also places some positive weight on the overseer's utility.

4 Dynamic Institutional Choice

We now turn to our institutional design problem, a process that is inherently dynamic–it depends on how the public forms its opinion about the regulator and on how this opinion evolves over time. We therefore build a fully dynamic model of institutional choice. We show that the static trade-off is present in the dynamic framework, but reputational dynamics add a novel element to the institutional choice problem: the unified institution uses the signal information efficiently for intervention decisions, but this slows down learning about the regulator's type. This dynamic trade-off can generate switches in institutional structure between unification and separation on the equilibrium path.

We present the result by focusing on the best relational contract for the public, i.e. the best sustainable equilibrium as introduced in Chari and Kehoe (1990). We argue this equilibrium concept captures the key dynamic interactions that motivate our model.

4.1 Dynamic Setup

In each period t, the incumbent regulator is endowed with the prior belief μ_t–the conditional probability that he is type H–given the public history. At the beginning of each period, the public P (acting through the government, which has the same objectives and information as the public) decides whether to continue with or to overhaul the regulatory agency. In case of an overhaul, a new regulator is selected from the pool of possible regulators, and the belief about R's type reverts to the initial prior given the distribution of types in the pool–the prior μ_H that the regulator is type H. The regulator's payoff each period is given by $u^R - c_\theta(e)$, where

$$u^R = \begin{cases} 1 & \text{if } R \text{ is kept} \\ 0 & \text{if } R \text{ is replaced through overhaul} \end{cases}.$$

Once a replacement happens, the removed regulator's continuation payoff is 0. The per-period payoff for O and P is equal to $u^O = u^P = -C$. The discount factor is common and equal to $\delta \in [0, 1)$.

The public then decides the institutional structure, $\psi \in \{0, 1\}$, where $\psi = 0$ denotes a separated structure and $\psi = 1$ a unified one. For simplicity, we assume that there is no switching cost to changing the institutional structure.[13] After the institutional structure is chosen, the intra-period play is the same as described in Section 2.

[13]If there were such a cost, then the institutional change would happen only if the benefit from the change were sufficiently large.

In order to focus on the trade-off between different institutional structures, we assume that (i) in the unified institution, the overseer O intervenes ($\iota = 1$) if and only if the signal is bad ($s_t = B$); and (ii) in the separated institution, the overseer does not intervene ($\iota = 0$). The interpretation is that the overseer has the statutory obligation to intervene if and only if the signal is bad. This obligation can be enforced, for example, through ex-post audits, which reveal the information available to the overseer at the time of the decision. Under this assumption, the overseer's action is automatic given ψ and s, so we can omit her from the strategic players.[14]

Throughout this section, we make the following additional assumption about the type L regulator:

Assumption 2 *The regulator of type L exerts only the minimal effort: $e_L = 0$.*

We make this assumption in order to focus on the optimal incentive structure for the regulator of type H. If the type L regulator were to exert additional effort, then the only difference would be the need to keep track of L's incentives. This analysis is relegated to Appendix A.3.

4.2 The Best Sustainable Equilibrium

We focus on the best sustainable equilibrium, which gives the best relational contract that the public can offer the regulator. To construct an equilibrium in this class, we consider a public randomization device for P, using the technique also employed in Ales et al. (2014). This randomization device is used to determine the replacement decision at the beginning of each period and the subsequent decisions within the period. At the beginning of each period t, the public draws a random variable $z_t \sim$ Uniform $[0, 1]$, which is observed to everyone.

Given z_t, the public decides whether or not to overhaul the regulatory agency. Let $p_t = 1$ denote replacing the current R through an overhaul, and $p_t = 0$ denote continuing with R.

[14]The analysis of the game would lead to the same equilibrium outcome if we assumed that O takes actions strategically to maximize her utility. More details are available upon request.

The public then decides the institutional structure $\psi_t \in \{0,1\}$. Afterwards, the regulator chooses effort e_t, and the overseer observes s_t if $\psi = 1$. The overseer intervenes ($\iota = 1$) and pays $C_t = C_1$ if and only if she observes $s_t = B$. Finally, if no intervention has happened, the overseer pays the final cost $C_t = C_2$ if the final outcome is bad ($y = B$). Otherwise, $C_t = 0$.

Since the public can only observe z_t, p_t, ψ_t, and C_t, the public outcome at the beginning of t is $h_t = (z_j, p_j, \psi_j, C_j)_{j=1}^{t-1}$. The public's strategy, σ_P, is a mapping from (h_t, z_t) to (p_t, ψ_t); the regulator's strategy, σ_R, is a mapping from $(h_t, z_t, p_t, \psi_t, \theta_t)$ to e_t, where θ_t is the private type of the regulator in period t. With the public randomization device, since the public and the regulator take actions sequentially and all the public actions are observable to the regulator, it is without loss of generality to focus on the pure strategy by the public. With a continuous action space $e_t \in [0,1]$ and a strictly convex cost function $c_H(e_t)$, the regulator always chooses the pure strategy. Given the pure strategy, we can assume without loss of generality that the strategy e_t does not depend on the past effort $(e_j)_{j=1}^{t-1}$.

Definition 3 *Given the history h_t of public outcomes up to period t, a sustainable equilibrium consists of continuation strategies $\{p_t(h_t, z_t), \psi_t(h_t, z_t)\}_t^\infty$ for the public, $\{e_t(h_t, z_t, p_t, \psi_t)\}_t^\infty$ for the regulator,*[15] *such that each player maximizes their respective expected continuation utility.*

As in Chari and Kehoe (1990), the on-path behavior of the best sustainable equilibrium for the public P can be characterized by the recursive method, subject to the constraint that there exists a punishment equilibrium which makes deviations unprofitable. We proceed to analyze the problem recursively, focusing on the best sustainable equilibrium for P. Afterwards, we show that there exists a punishment equilibrium that sustains this on-path behavior.

In particular, we study the problem of maximizing P's welfare given a promised payoff V to the type H regulator.[16] Given our motivation, this structure resembles a feasible reward

[15] Since type L exerts zero effort, we omit $\theta_t = H$ from the regulator's strategy.
[16] Since type L does not work, we do not keep track of his utility.

structure for the regulator. It allows P to write a sustainable contract for R that promises a certain level of total benefits to R. We will verify that the public has the incentive to fulfill this promise given the threat of punishment. In addition, given the public history h_t, the regulator's strategy σ_R, and the fact that the probability of a regulator of type H is μ_H upon the initial appointment, the public can calculate the conditional probability that the current regulator is type H, denoted as the belief μ.

Let $J(\mu, V)$ be P's welfare given these two state variables, the belief μ and the promised utility V. Upon an overhaul, a replacement occurs and the belief goes back to μ_H, but P can choose the promised utility to the new regulator to maximize its welfare. Hence, P's utility after replacing R, denoted by $\bar{J}$, should satisfy

$$\bar{J} = \max_V J(\mu_H, V).$$

The problem for P is to select a vector $\alpha_z = (p_z, \psi_z, e_z, V_z')$ for each possible realization of z, where $p_z \in \{0, 1\}$ is the index for the replacement, with $p_t = 1$ denoting replacement, ψ_z is the institutional structure of that period, e_z is the proposed effort for the type H regulator,[17] and $V_z'(o_z)$ is the promised continuation value for the regulator of type H given the public outcome o_z in the current period. Here, the public outcome given z, denoted by o_z, consists of ψ_z, ι_z, and C. Since the on-path outcome has full support, the public always believes that the recommended effort e_z is taken. Hence the next period's belief $\mu'(\mu, e_z, o_z)$ is determined by Bayes' rule, which depends on the prior μ, the recommended effort e_z, and the public outcome o_z.

The public P chooses α_z to solve the following dynamic program. For each $\mu \in [0, 1]$ and $V \in [0, 1]$,

$$\begin{aligned} J(\mu, V) &= \max_\alpha \int_z \Big[p_z \bar{J} + (1 - p_z)\Big\{(1-\delta)\, u^P(\mu, \psi_z, e_z) \\ &\qquad + \delta \sum_{o_z} \Pr(o_z | \mu, \psi_z, e_z)\, J\left(\mu'(\mu, e_z, o_z), V_z'(o_z)\right)\Big\}\Big]\, dz, \quad (14) \end{aligned}$$

[17] The L-type always chooses $e_L = 0$.

where

$$u^P(\mu, \psi_z, e_z) = -\mu \mathbb{E}[C|\psi_z, e_z] - (1-\mu)\mathbb{E}[C|\psi_z, e_L = 0]$$

is the public's instantaneous utility. With probability μ, the regulator is type H and chooses e_z and with the remaining probability, he is type L and takes $e = 0$.

The public is subject to the following constraints:

$$V = \int_z (1-p_z)\left\{(1-\delta)[1-c_H(e_z)] + \delta \sum_{o_z} \Pr(o_z|\psi_z, e_z) V'_z(o_z)\right\} dz; \tag{15}$$

$$e_z \in \arg\max \left\{(1-\delta)[1-c_H(e_z)] + \delta \sum_{o_z} \Pr(o_z|\psi_z, e_z) V'_z(o_z)\right\}; \tag{16}$$

$$V'_z(o_z) \in [0,1] \text{ for each } o_z. \tag{17}$$

Constraint (15) is the promise keeping that P is bound to in equilibrium, and (16) is the incentive compatibility constraint for R.[18] Constraint (17) places the upper and lower bounds on the future promised value since (i) 0 is the minimum payoff that the regulator receives; and (ii) 1 is the highest feasible utility for the regulator, which is implemented by keeping the regulator and allowing him put in effort $e = 0$ forever.[19]

The next lemma shows that the above maximization problem characterizes the best sustainable equilibrium for P.

Lemma 1 *The best sustainable equilibrium for P is characterized by the mapping from $(\mu, V) \in [0,1] \times [0,1]$ to a vector $(p_z, \psi_z, e_z, \iota_z, V'_z)_{z\in[0,1]}$ which maximizes (14) subject to (15), (16), and (17).*

Proof. In Appendix A.1.4. ■

By definition, objective (14) gives the highest welfare for P given the promise keeping and incentive compatibility constraints. Hence, we are left to construct a punishment equilibrium

[18] Notice that, from P's view point, the distribution of the outcome $\Pr(o_z|\mu, e_z)$ depends on μ, while for the promise keeping and incentive compatibility, we have $\Pr(o|e_z)$ since R knows his own type.

[19] This is also why we define the dynamic program only for $V \in [0,1]$.

to sustain this outcome on the equilibrium path. Note that there is always an equilibrium in which R exerts no effort, expecting replacement every period; given this, P picks the unified institution and replaces the regulator every period. In the sustainable equilibrium, the lowest value P could obtain would be $J(\mu, 1)$, since the promised value 1 allows R to stay without putting in effort. This value corresponds to the "no effort equilibrium." Then, switching to this worst equilibrium after any deviation from policy α_z, we can sustain the solution for (14) on the equilibrium path.

4.3 Equilibrium Properties

We first show that P's value function $J(\mu, V)$ is concave in V, convex and increasing in μ:

Lemma 2 *$J(\mu, V)$ is concave is V, convex in μ, and increasing in μ (strictly increasing if $V \in (0, 1)$).*

Proof. In Appendix A.1.5. ■

Since we allow the public randomization, the concavity of $J(\mu, V)$ with respect to V follows from the standard arguments. The convexity of $J(\mu, V)$ with respect to μ follows from the following exercise. Assume that P receives additional information to update its belief about R's type. The belief is martingale, so the updated belief is a mean-preserving spread of the original belief μ. Since P can always ignore this new information, this mean-preserving spread is always (at least weakly) welfare-improving. Finally, $J(\mu, V)$ is increasing in μ since P can always choose the same continuation strategy when the belief is $\mu' > \mu$ as it does when the belief is μ. Since type L chooses $e_L = 0$, this allows P to obtain at least weakly higher welfare with μ' compared to μ.

We can then derive the following implications about the shape of $J(\mu, V)$.

Lemma 3 *The value function $J(\mu, V)$ has the following properties:*

1. $J(\mu, 0) = \bar{J}\ \forall \mu$;

2. *For each μ, there exists $V(\mu) \in \mathbb{R}$ such that $J(\mu, V)$ is linear for $V \in [0, V(\mu)]$, where $V(\mu) \geq 1 - \delta$, with strict inequality for $\mu \geq \mu_H$. Moreover, the slope for the linear part, $\frac{d}{dV} J(\mu, V)\big|_{V \in [0, V(\mu)]}$, is negative for $\mu < \mu_H$; zero for $\mu = \mu_H$; and positive for $\mu > \mu_H$.*

3. *There exists $V^*(\mu) \in \arg\max_V J(\mu, V)$. For each μ, at $V = V^*(\mu)$, the following properties hold:*

   (a) $V_z'(o_z) \leq \arg\max_V J(\mu'(\mu, e_z, o_z), V)$ for each event o_z with negative belief update $\mu'(\mu, e_z, o_z) \leq \mu$;

   (b) $V_z'(o_z) \geq \arg\max_V J(\mu'(\mu, e_z, o_z), V)$ for each event o_z with positive belief update $\mu'(\mu, e_z, o_z) \geq \mu$.

Proof. In Appendix A.1.6. ■

The above properties have the following intuition. First, if the value promised to R is 0, only immediate replacement can fulfill this promise. Hence, it must be the case that $J(\mu, 0) = \bar{J}$ for each μ. Second, for a sufficiently small promised value V, replacement must happen with positive probability–otherwise R would obtain a higher payoff than promised. This positive probability is generated by randomizing between keeping and replacing R. We can show that it is optimal for P to promise the same utility $V(\mu)$ if R is kept. Then, any $V \in [0, V(\mu)]$ can be generated by varying the probability of keeping versus replacing R.

The third property follows from concavity: the function $J(\mu, V)$ has a maximum, denoted $V^*(\mu)$. Consider the case in which at $V = V^*(\mu)$ the outcome o_z is such that the belief is updated negatively, $\mu' < \mu$. The negative belief update implies that o_z is the outcome which happens less often with higher effort. Hence, reducing $V_z'(o_z)$ incentivizes R to supply more effort. This increased effort has two benefits: the instantaneous welfare u^P is improved; and, since higher effort makes type H more distinguishable from type L, the belief update is larger. As long as $V_z'(o_z) > \arg\max_V J(\mu'(\mu, e_z, o_z), V)$, reducing $V_z'(o_z)$ directly improves P's continuation welfare as well; however, the opposite happens if $V_z'(o_z) < \arg\max_V J(\mu'(\mu, e_z, o_z), V)$. Intuitively, promising too little continuation payoff to R forces P to replace R even if P believes that the regulator is type H. The analysis when outcome o_z leads to a positive belief update is analogous.

Finally, the second property also implies that $\arg\max_V J(\mu', V) = 0$ if the updated belief μ' becomes less than the prior in the pool, μ_H, right after a new regulator is selected. This is consistent with the utility specification (9) we assumed in the static model.

4.4 The Optimal Dynamic Policy

We now solve for the optimal dynamic policy. For the following set of results in the analysis, we make a simplifying assumption regarding the error in the signal s.

Assumption 3 $\Pr(s = B|y = G, e) = \varepsilon > 0$ *and* $\Pr(s = G|y = B, e) = 0$.

Under this assumption, we consider the case in which the error happens in the signal only when the final outcome is good. This makes the unified institution even better compared to the case in which there is also error after the bad outcome. Hence, the benefit of the separated institution becomes even stronger without this assumption.

Given this simplification of the signal error, the outcome $o_z = C_2$ never happens when $\psi_z = 1$, since there is no error associated with $s = B$, and the overseer responds to this signal by intervening. Using these simplified public outcomes, we can pin down the continuation utility for R after each event.

Lemma 4 *The continuation utility for R is higher after outcome $o_z = 0$ than after outcome $o_z \in \{C_1, C_2\}$. Moreover, the difference between R's expected continuation payoff after $o_z = 0$ and the continuation payoff after $o_z \in \{C_1, C_2\}$ increases as a function of effort e_z.*

Proof. In Appendix A.1.7. ■

The intuition for this result is that R is rewarded with a higher promised payoff after the outcomes that indicate higher effort. Conversely, R is punished with lower promised payoff after the outcome that indicates low effort. The difference between the reward and the punishment must be larger when P implements higher effort.

Comparing the difference between the reward and punishment offered to R under each institutional structure, we obtain the following result.

Lemma 5 *Fixing any targeted effort level e_z, the continuation utility for R satisfies*

$$V_z'(0|\psi=1) - V_z'(C_1|\psi=1) > V_z'(0|\psi=0) - V_z'(C_2|\psi=0).$$

Proof. In Appendix A.1.7. ■

To implement any effort e_z, the public must incentivize R with a higher relative reward under the unified institution compared to the separated institutional structure. The intuition for this result is the following. Under the unified institution, the observed outcome is less informative about R's effort, because the intervention happens after $s = B$ even if $y = G$. Effort is then less conducive to higher continuation payoff, so R is less willing to exert high effort unless he is promised a higher reward.

The results so far show how the static trade-off translates to the dynamic framework. For any effort level e, P's instantaneous utility is higher in the unified institutional structure, since O's use of the signal information lowers the expected cost; however, in order to attain the same level of effort in the unified institution as in the separated structure, P has to promise a higher reward to R after the good outcome. This in turn reduces P's future payoff, since P essentially accepts less effort in the future. Formally, part 3 of Lemma 3 shows that, if we start from the optimal promised value given the current belief, then increasing the continuation payoff for R after the good outcome reduces the continuation utility for P. This corresponds to the static trade-off of ex-ante high-powered incentives versus ex-post efficient use of the information.

The following result highlights the new trade-off that emerges when dynamics are considered.

Proposition 4 *Fixing any targeted effort level e_z, the following dynamic trade-off emerges: under institutional separation ($\psi = 0$) learning about R's type happens at a faster pace–the update in beliefs each period is larger.*

Proof. In Appendix A.1.8. ■

Consider the update in beliefs after the worst outcome in each institutional structure in equilibrium, i.e., $o_z = C_1$ in the unified institution and $o_z = C_2$ in the separated institution. It then follows from Bayes' rule that

$$\underbrace{\left|\mu - \frac{\mu\left[1-(1-\varepsilon)\,q\left(e_z\right)\right]}{1-(1-\varepsilon)\left[\mu q\left(e_z\right)+(1-\mu)\,q\left(0\right)\right]}\right|}_{\substack{\mu'(C_1|\psi=1):\ \text{belief update} \\ \text{with the unified institution}}} < \underbrace{\left|\mu - \frac{\mu\left[1-q\left(e_z\right)\right]}{1-\left[\mu q\left(e_z\right)+(1-\mu)\,q\left(0\right)\right]}\right|}_{\substack{\mu'(C_2|\psi=0):\ \text{belief update} \\ \text{with separate institutions}}}, \tag{18}$$

so the beliefs are updated more strongly under the separated institutional structure. The reason is that the final bad outcome $y = B$ is more informative of R's effort than the signal $s = B$. The same calculation shows that the belief update is the same between $\psi = 1$ and $\psi = 0$ after $y = G$. Since P's belief is a martingale, this means that the updated belief distribution when $\psi = 0$ is a mean-preserving spread of the distribution when $\psi = 1$. This difference in the speed of learning translates into P's utility. By Lemma 2, since $V(\mu, V)$ is convex in μ, faster learning increases P's utility. Therefore, faster learning under separated institutions increases the benefit to P.

We can show that the dynamic advantage brought about by institutional separation is highest for intermediate values of the belief μ.

Corollary 2 *Fixing any targeted effort level e_z, the difference in learning pace between separation and unification is largest when beliefs about R being type H are intermediate and smallest when beliefs are either* 0 *or* 1.

Proof. In Appendix A.1.8. ■

The result follows from examining the dependence of the updated belief μ' on the prior μ. The intuition is that the advantage of learning fast is largest when there is enough uncertainty about R's type.

Finally, the interplay of the static and the dynamic trade-offs described above opens up the possibility of switches between institutional unification and separation.

Proposition 5 *Switches in institutional structure, from unification to separation and from separation to unification, can emerge on the equilibrium path.*

This conclusion follows because the advantage of institutional separation is highest when the dynamic trade-off is strongest in favor fast learning–when the beliefs about R are intermediate. When P believes that R is type L with higher probability, the gains from learning are low. As more good outcomes are observed, P increases its belief that the regulator is type H, and the gains from learning increase. Finally, with enough good outcomes, P becomes almost certain that R is type H and new observations bring a low contribution to learning. This dynamic leads to a choice of institutional unification when R's reputation is low, the switch to separation when R's reputation is intermediate, and then a switch back to unification again when R's reputation is sufficiently high.

This dynamic can explain the fluctuations we see in the data presented in Figure 1 between unification and separation, as well as the different institutional choices made by countries during the same time period.

4.5 Numerical Illustration

To illustrate how the change in institutional structure crucially depends on the dynamics of learning, we provide the following numerical example. Consider the case with $\delta/(1-\delta) = 0.89$, $C_2/C_1 = 1.56$, $\varepsilon = 0.2$ and

$$q(e) = 0.2 + 0.7e;\ c_H(e) = \frac{e^2}{10(1-e)},$$

As assumed above, $e_L = 0$.

We begin by computing the social welfare in the static game described in Section 3.[20] Figure 2 graphs the social welfare against the prior $\mu_H \in [0,1]$ for the two institutional

[20]Here, to make the static and dynamic models comparable, we assume $\Pr(s = B|y = G) = \varepsilon$ and $\Pr(s = G|y = B) = 0$ in the static model as well. All the qualitative results still hold without this assumption.

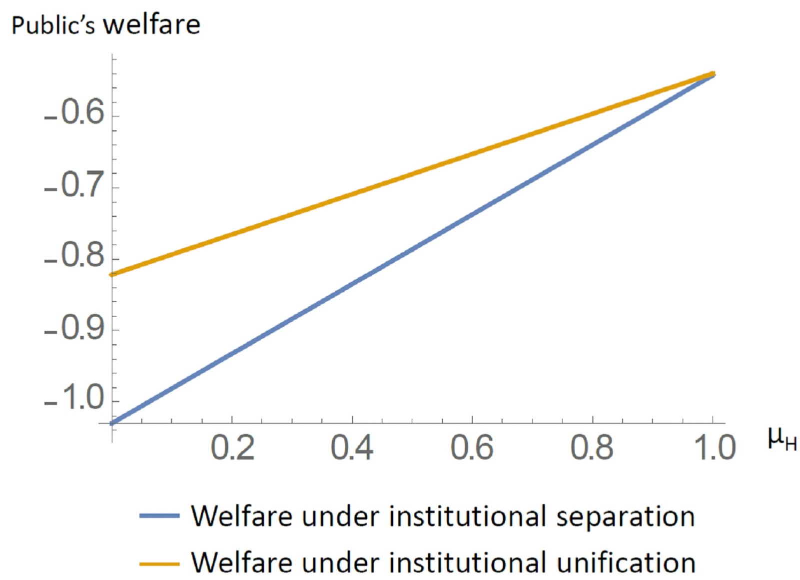

Figure 2: Public's welfare in the static game

setups. In the numerical example we chose, the public's welfare in the static game is higher under institutional unification for all possible priors μ_H.

Next, we compute the solution to the dynamic game taking $\mu_H = 0.4$, the prior that the regulator is of type H in the initial distribution. We consider the case in which the good outcome is realized for 20 consecutive periods. The upper panels of Figure 3 show the evolution of the promised value V to the regulator and the equilibrium regulatory effort provided by type H. In response to each observation of a good outcome, the public optimally increases the value promised to the regulator. Since this implies a lower probability of removal in the future periods, the regulator optimally reduces effort.

The lower left-hand side panel of Figure 3 shows the evolution of the regulator's reputation. After observing each good outcome, the public positively updates their belief about the regulator's type, until it becomes certain that he is type H.

Finally, the lower right-hand side panel of Figure 3 displays the equilibrium institutional choice in the best sustainable equilibrium. The equilibrium institutional structure at $\mu = \mu_H$ is unification. Since $\mu_H < 0.5$, as good outcomes accumulate, the public's belief about the regulator increases towards 0.5. Then, the dynamic benefit of faster learning becomes more

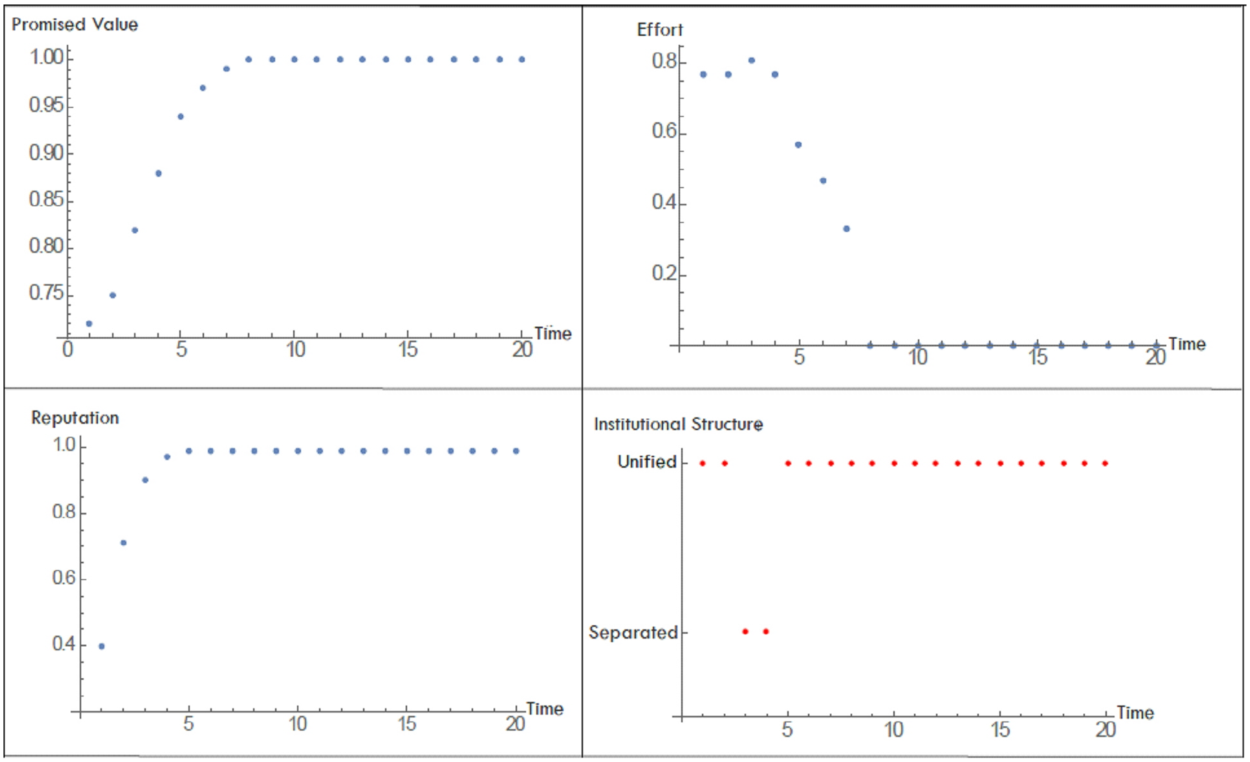

Figure 3: Outcomes in the dynamic game when the good outcome is realized for 20 periods

important, which leads to separation being optimal. As more good outcomes arrive, the belief about the regulator becomes sufficiently high, so the value of learning decreases. This decreases the benefit of separation. Also, as the regulator is promised a higher value, the effort provision decreases. In the limit, this leads to the static trade-off vanishing, since effort becomes minimal. Therefore, the public optimally changes the institutional structure to unification, to take advantage of the availability of information to the overseer.

5 Conclusion

This paper considered a key question in the design of regulatory institutions: how to optimally structure regulatory institutions between which there exists an oversight relationship, when information flows are important. We examined the choice between institutional unification and separation, where unification affords the overseer greater access to information.

This creates a static trade-off between effort provision by the regulator and access to information by the overseer. Dynamically, another force emerges: when the overseer has less access to information, the government is better able to learn about the regulator's quality. The interaction of these forces can generate switches between unification and separation on the equilibrium path.

The present model provides a framework for exploring the links between regulatory structure, public perception of institutional quality, and economic outcomes. The dynamics of the model imply a link between public perceptions of institutional quality and the timing of structural regulatory reforms. This link can be studied both within countries and across countries, given the wide variation in government responses to similar crises. For the case of financial regulation, the static trade-off between effort and information implies that separation of bank supervision from the lender of last resort should provide incentives for higher supervisory effort. This effort should translate into fewer crisis episodes. Some suggestive evidence to this end is provided in Eichengreen and Dincer (2011), who look at bank performance and supervisory structure in 140 countries from 1998 to 2010. They find that countries with separated institutions have fewer nonperforming loans as a share of GDP. Our model suggests that these results should extend to banking crisis episodes, not only nonperforming loans. Further empirical examination of the model's predictions can shed light on the magnitude of these forces and their effect on regulatory outcomes.

References

Aghion, P., Tirole, J., 1997. Formal and real authority in organizations. Journal of political economy, 1–29.

Ales, L., Maziero, P., Yared, P., 2014. A theory of political and economic cycles. Journal of Economic Theory 153, 224–251.

Alesina, A., Tabellini, G., 2007. Bureaucrats or politicians? part i: a single policy task. The American Economic Review 97 (1), 169–179.

Banks, J. S., Sundaram, R. K., 1993. Adverse selection and moral hazard in a repeated elections model. ch 12, 295–311.

Banks, J. S., Sundaram, R. K., 1998. Optimal retention in agency problems. Journal of Economic Theory 82 (2), 293–323.

Bardhan, P., Mookherjee, D., 2006. Decentralization, corruption, and government accountability. International handbook on the economics of corruption 6, 161–188.

Besley, T., Coate, S., 2003a. Centralized versus decentralized provision of local public goods: a political economy approach. Journal of public economics 87 (12), 2611–2637.

Besley, T., Coate, S., 2003b. Elected versus appointed regulators: Theory and evidence. Journal of the European Economic Association 1 (5), 1176–1206.

Blanchard, O., Shleifer, A., 2001. Federalism with and without political centralization: China versus russia. IMF Economic Review 48 (1), 171–179.

Buisseret, P., 2016. Together or apart? on joint versus separate electoral accountability. The Journal of Politics 78 (2), 542–556.

Cai, H., Treisman, D., 2005. Does competition for capital discipline governments? decentralization, globalization, and public policy. The American Economic Review 95 (3), 817–830.

Carpenter, D., 2010. Reputation and Power. Princeton University Press.

Carpenter, D. P., 2001. The Forging of Bureaucratic Autonomy. Princeton University Press.

Chari, V. V., Kehoe, P. J., 1990. Sustainable plans. Journal of political economy, 783–802.

De Mesquita, E. B., Stephenson, M. C., 2007. Regulatory quality under imperfect oversight. American Political Science Review 101 (03), 605–620.

Dovis, A., Golosov, M., Shourideh, A., 2016. Political economy of sovereign debt: A theory of cycles of populism and austerity. NBER Working Paper.

Eichengreen, B., Dincer, N., 2011. Who should supervise? the structure of bank supervision and the performance of the financial system. NBER Working Paper (w17401).

Fox, J., Van Weelden, R., 2010. Partisanship and the effectiveness of oversight. Journal of Public Economics 94 (9), 674–687.

Hayek, F. A., 1945. The use of knowledge in society. The American economic review 35 (4), 519–530.

Levitt, S. D., Snyder, C. M., 1997. Is no news bad news? information transmission and the role of" early warning" in the principal-agent model. The RAND Journal of Economics, 641–661.

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

Melecky, M., Podpiera, A. M., 2013. Institutional structures of financial sector supervision, their drivers and historical benchmarks. Journal of Financial Stability 9 (3), 428–444.

Oates, W. E., 1972. Fiscal federalism. Harcourt Brace Jovanovich Academic Press.

Persson, T., Roland, G., Tabellini, G., 1997. Separation of powers and political accountability. The Quarterly Journal of Economics, 1163–1202.

Prat, A., 2005. The wrong kind of transparency. American Economic Review 95 (3), 862–877.

Stokey, N. L., 1989. Recursive methods in economic dynamics. Harvard University Press.

Tiebout, C. M., 1956. A pure theory of local expenditures. The journal of political economy, 416–424.

Weingast, B. R., 1995. The economic role of political institutions: Market-preserving federalism and economic development. Journal of Law, Economics, & Organization, 1–31.

A Appendix

A.1 Proofs

A.1.1 Proof of Proposition 1

Part 1. Effort strategies

It is sufficient to show that e_θ is weakly increasing in θ, since if $e_H = e_L$, then $b|C = \mu$ for each outcome C, and each regulator puts in zero effort.

Let $v(\theta)$ be the equilibrium payoff of R with type $\theta \in \{H, L\}$. We compare types H and L. When R with type H chooses the same effort e_L as type L, he induces the same distribution over reputation as L. For such a deviation not to be profitable, we have

$$v(H) \geq v(L) + c_L(e_L) - c_H(e_L).$$

Note that $-(c_H(e_L) - c_L(e_L)) = c_L(e_L) - c_H(e_L)$ is the reduction in the cost that type H has to pay to exert effort e_L compared to type L. Similarly, the condition for R with L to not want to pretend to be type H, we have

$$v(L) \geq v(H) + c_H(e_H) - c_L(e_H).$$

In total, we have

$$c_L(e_H) - c_H(e_H) \geq v(H) - v(L) \geq c_L(e_L) - c_H(e_L).$$

Rearranging these inequalities yields

$$c_L(e_H) - c_L(e_L) \geq c_H(e_H) - c_H(e_L).$$

Since we assumed that the marginal cost of e is decreasing in L, we have $e_H \geq e_L$, as desired.

Part 2. Intervention strategy

By Assumption 1, for any effort level of the regulator, without additional information, it is optimal not to intervene.

Part 3. Beliefs

We focus on the equilibrium with $e_H > e_L$. By Bayes' rule, we have

$$b|0 = \frac{\mu q(e_H)}{\mu q(e_H) + (1-\mu) q(e_L)} > \mu$$

and

$$b|C_2 = \frac{\mu (1 - q(e_H))}{\mu (1-q(e_H)) + (1-\mu)(1-q(e_L))} < \mu$$

since $q'(e_H) > 0$ (otherwise $e_H > 0$ is not optimal given $c_H(e) > 0$ for each $e > 0$).

A.1.2 Proof of Proposition 2

Part 1. Effort strategies

The proof that $e_H \geq e_L$, with strict inequality if $e_H > 0$, is the same as the one in the proof of Proposition 1, Part 1.

Part 2. Intervention strategy

By Assumption 1, for any effort level of the regulator, it is optimal to intervene if and only if $s = B$.

Part 3. Beliefs

By Bayes' rule,

$$\begin{aligned}
b|0 &= \frac{\mu q(e_H)(1-\varepsilon)}{\mu q(e_H)(1-\varepsilon) + (1-\mu) q(e_L)(1-\varepsilon)} = \frac{\mu q(e_H)}{\mu q(e_H) + (1-\mu) q(e_L)}, \\
b|C_1 &= \frac{\mu [q(e_H)\varepsilon + (1-q(e_H))(1-\varepsilon)]}{\mu [q(e_H)\varepsilon + (1-q(e_H))(1-\varepsilon)] + (1-\mu)[q(e_L)\varepsilon + (1-q(e_L))(1-\varepsilon)]}, \\
b|C_2 &= \frac{\mu (1-q(e_H))\varepsilon}{\mu (1-q(e_H))\varepsilon + (1-\mu)(1-q(e_L))\varepsilon} = \frac{\mu (1-q(e_H))}{\mu (1-q(e_H)) + (1-\mu)(1-q(e_L))}.
\end{aligned}$$

Since we focus on the case with $e_H > e_L$ and $q(e)$ is increasing, we have

$$b|0 > \mu > b|C_1 > b|C_2.$$

A.1.3 Proof of Proposition 3

Dividing (13) by C_1, the unified institution is more costly if and only if

$$\begin{aligned} V(\varepsilon, c) \equiv & -\sum_{\theta} \Pr(\theta)\left(1-q\left(e_{\theta}^{s}\right)\right) c \\ & +\sum_{\theta}\left\{\Pr(\theta)\left((1-\varepsilon)\left(1-q\left(e_{\theta}^{\varepsilon}\right)\right)+\varepsilon q\left(e_{\theta}^{\varepsilon}\right)\right)\right. \\ & \left.+\Pr(\theta) \varepsilon\left(1-q\left(e_{\theta}^{\varepsilon}\right)\right) c\right\}>0, \end{aligned}$$

where $c=\frac{C_2}{C_1}$. Here, $e_\theta^s = e_\theta^0$, since (10) and (11) imply that the effort level under separated institutions corresponds to that under the unified institution when $\varepsilon = 0$.

Taking the total derivative, we obtain

$$\begin{aligned} & \sum_{\theta} \Pr(\theta)\left[\begin{array}{c} q\left(e_{\theta}^{\varepsilon}\right)+\left(1-q\left(e_{\theta}^{\varepsilon}\right)\right)(c-1) \\ -(1-2 \varepsilon) q^{\prime}\left(e_{\theta}^{\varepsilon}\right) \frac{d e_{\theta}^{\varepsilon}}{d \varepsilon}-\varepsilon q^{\prime}\left(e_{\theta}^{\varepsilon}\right) \frac{d e_{\theta}^{\varepsilon}}{d \varepsilon} c \end{array}\right] d \varepsilon \\ & +\sum_{\theta} \Pr(\theta)\left[\varepsilon\left(1-q\left(e_{\theta}^{\varepsilon}\right)\right)-\left(1-q\left(e_{\theta}^{0}\right)\right)\right] d c . \end{aligned}$$

By (12), we obtain $\frac{de_\theta^\varepsilon}{d\varepsilon} < 0$. Since $c > 1$ and $\varepsilon < \frac{1}{2}$, the coefficient of $d\varepsilon$ is positive. Moreover, as long as $V(\varepsilon, c) \leq 0$, we have

$$\varepsilon\left(1-q\left(e_{\theta}^{\varepsilon}\right)\right)-\left(1-q\left(e_{\theta}^{0}\right)\right)<0.$$

Hence $V(\varepsilon, c)$ is increasing in ε and decreasing in c, as desired.

Finally, to show that we can have $V(\varepsilon, c) > 0$ while satisfying Assumption 1, we create the following example:

$$\begin{aligned} q(e) &= 0.4+0.3 e, \\ c_{\theta}(e) &= e^{3/2}\left[m_{\theta}(1-e)\right]^{-1}, \end{aligned}$$

with $m_H = 10, m_L = 0.8$. Also, $\mu^H = 0.97$.

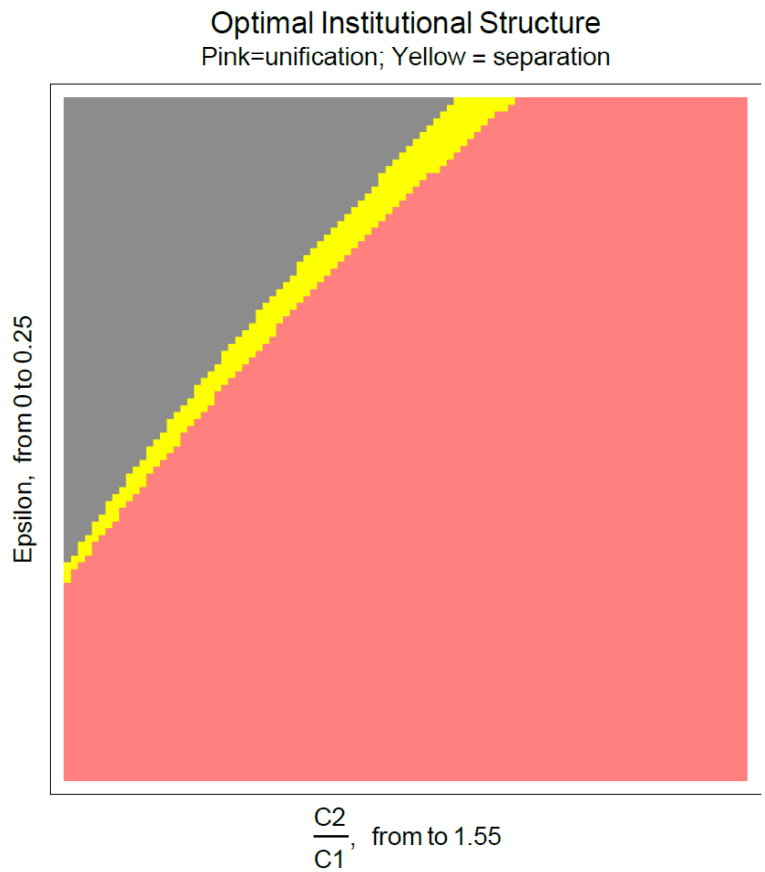

Figure 4: Optimal institutional structure in the numerical example

Notice that these functional forms imply that a regulator will choose an effort level at most $e_{\max}$, which is a solution for

$$q'(e) = c'_H(e),$$

since the most high-powered incentive is provided when the regulator receives the utility of 0 if and only if the final bad outcome is reached. Taking this into account, we can modify Assumption 1 to

$$1 + \frac{q(e_{\max})}{1 - q(e_{\max})} \frac{\varepsilon}{1 - \varepsilon} < \frac{C_2}{C_1} < 1 + \frac{q(0)}{1 - q(0)} \frac{1 - \varepsilon}{\varepsilon} \tag{19}$$

and

$$\frac{1}{1 - q(0)} > \frac{C_2}{C_1}. \tag{20}$$

We consider $\varepsilon \in [0, 0.25]$ and $C_2/C_1 \in [1, 2]$.

Figure 4 shows the optimal choice of institutional structure for $(\varepsilon, C_2/C_1)$ in $[0, 0.25] \times [1, 2]$. The grey area marks the values $(\varepsilon, C_2/C_1)$ at which conditions (19) or (20) are not satisfied. The pink area denotes the region in which unification is optimal, and the yellow area denotes the region in which separation is optimal. Figure 4 illustrates a case in which (i) there exist parameter values that satisfy (19) and (20); and (ii) separation is optimal for sufficiently high ε and sufficiently low C_2/C_1.

A.1.4 Proof of Lemma 1

Let $\underline{SW}$ be P's utility under the no effort equilibrium. In this equilibrium, after each public history, P replaces R with probability one, and R exerts zero effort. This is a sequential equilibrium since (i) it is optimal for P to replace the regulator given that the future regulator supplies zero effort and (ii) it is optimal for the regulator to supply zero effort since he is replaced for sure.

By feasibility, $V \in [0, 1]$ since 1 is the maximum average payoff that the public can deliver by letting $p_z = e_z = 0$ for each z. If $V = 1$, since R does not work and is kept forever, so we have $J(\mu, 1) = \underline{SW}$ for each μ. On the other hand, if $V = 0$, then P replaces R right away, and so $J(\mu, 0) = \bar{J} \geq \underline{SW}$ for each μ. Since $J(\mu, V)$ is concave in V by Lemma 2,[21] for each μ and $V \in [0, 1]$, we have $J(\mu, V) \geq \underline{SW}$.

Consider the following strategy: On the equilibrium path, the public history h^t decides μ and V. Given z, as long as the public chooses p_z and ψ_z corresponding to the solution to the dynamic program, the regulator supplies e_z. The regulator's deviation is ignored. If P deviates from this equilibrium path, then P chooses $p_z = 1$ and $\psi_z = 1$ forever, and R chooses $e_z = 0$ forever (switching to the no effort equilibrium). Incentive compatibility (16) ensures the regulator's incentive, and $J(\mu, V) \geq \underline{SW}$ ensures the public's incentive.

It is useful to verify that the welfare at the arrival of a new regulator is higher than this no effort equilibrium:

$$\bar{J} > \underline{SW}. \tag{21}$$

[21] Note that proof of Lemma 2 does not depend on Lemma 1.

To see why, the public can improve upon $\underline{SW}$ as follows: The public always takes $\psi_z = 1$ as in the no effort equilibrium. If $o_z = C_2$, then P keeps the regulator forever. Otherwise, P replaces the regulator (and goes back to the no effort equilibrium). That is, P rewards the regulator after a good outcome in the first period, which incentivizes the regulator to supply positive effort. Hence, the public can obtain a payoff greater than $\underline{SW}$ in the first period, and then go back to the no effort equilibrium. In total, P obtains a payoff higher than $\underline{SW}$.

A.1.5 Proof of Lemma 2

Part 1. Concavity with respect to V.

We show that $J(\mu, V)$ is concave in V, for a fixed μ. Suppose $V = \beta V_1 + (1-\beta) V_2$ for $V_1, V_2, \beta \in [0,1]$; and let $\alpha[V_1]$ and $\alpha[V_2]$ be the optimal policies for (μ, V_1) and (μ, V_2), respectively.

Suppose P chooses $\alpha[V_1]$ with probability β and $\alpha[V_2]$ with probability $1-\beta$, according to the realization of the public randomization device. We need to check that (i) P's choice satisfies the promise keeping and incentive compatibility constraints; and P's value is at least $\beta J(\mu, V_1) + (1-\beta) J(\mu, V_2)$.

1. Since both $\alpha[V_1]$ and $\alpha[V_2]$ satisfy promise keeping

$$\begin{aligned} V_1 &= \int_{z_1} [(1 - p_{z_1}[V_1]) \{(1-\delta)[1 - c_H(e_{z_1}[V_1])] \\ &\quad + \delta \sum_{o_z} \Pr\left(o_z | \psi_{z_1}[V_1], e_{z_1}[V_1]\right) V_z'[V_1](o_z)\}] dz_1 \end{aligned}$$

and

$$\begin{aligned} V_2 &= \int_{z_2} [(1 - p_{z_2}[V_2]) \{(1-\delta)[1 - c_H(e_{z_2}[V_2])] \\ &\quad + \delta \sum_{o_z} \Pr\left(o_z | \psi_{z_2}[V_2], e_{z_2}[V_2]\right) V_z'[V_2](o_z)\}] dz_2, \end{aligned}$$

we have

$$
\begin{aligned}
V &= \beta V_1 + (1-\beta) V_2 \\
&= \beta \int_{z_1} (1 - p_{z_1}[V_1]) \{(1-\delta)[1 - c_H(e_{z_1}[V_1])] \\
&\quad + \delta \sum_{o_z} \Pr\left(o_z | \psi_{z_1}[V_1], e_{z_1}[V_1]\right) V_z'[V_1](o_z)\} dz_1 \\
&\quad + (1-\beta) \int_{z_2} (1 - p_{z_2}[V_2]) \{(1-\delta)[1 - c_H(e_{z_2}[V_2])] \\
&\quad + \delta \sum_{o_z} \Pr\left(o_z | \psi_{z_2}[V_2], e_{z_2}[V_2]\right) V_z'[V_2](o_z)\} dz_2.
\end{aligned}
$$

Hence, promise keeping is satisfied.

2. Conditional on the realization of the public randomization device, since both $\alpha[V_1]$ and $\alpha[V_2]$ are incentive compatible, R with type H will choose $e[V_1]$ and $e[V_2]$ according to the realization of the public randomization device. Hence, incentive compatibility is satisfied.

Therefore, we are left to verify that this policy gives P at least the value $\beta J(\mu, V_1) + (1-\beta) J(\mu, V_2)$. With probability β, we achieve $J(\mu, V_1)$ and with probability $1-\beta$, we achieve $J(\mu, V_2)$ since we fix μ. Hence we achieve $\beta J(\mu, V_1) + (1-\beta) J(\mu, V_2)$, as desired.

Part 2. Convexity with respect to μ.

Let $J(\mu, V, \theta)$ be the social welfare when the public follows the optimal strategy given (μ, V), and the current type is $\theta \in \{H, L\}$. We have

$$
\begin{aligned}
J(\mu, V) &= \mu J(\mu, V, H) + (1-\mu) J(\mu, V, L) \\
&= J(\mu, V, L) + \mu [J(\mu, V, H) - J(\mu, V, L)].
\end{aligned}
$$

Take μ, μ_1, μ_2 and $\beta \in [0,1]$ such that $\mu = \beta\mu_1 + (1-\beta)\mu_2$. By taking the strategy given

(μ, V) when the belief is μ_1, P obtains

$$\begin{aligned} & \mu_1 J(\mu, V, H) + (1-\mu_1) J(\mu, V, L) \\ = \; & J(\mu, V, L) + \mu_1 [J(\mu, V, H) - J(\mu, V, L)] \\ \leq \; & J(\mu_1, V). \end{aligned}$$

Similarly,

$$J(\mu, V, L) + \mu_2 [J(\mu, V, H) - J(\mu, V, L)] \leq J(\mu_2, V).$$

Hence,

$$\begin{aligned} & \beta J(\mu_1, V) + (1-\beta) J(\mu_2, V) \\ \geq \; & \beta J(\mu, V, L) + \beta \mu_1 [J(\mu, V, H) - J(\mu, V, L)] \\ & + (1-\beta) J(\mu, V, L) + (1-\beta) \mu_2 [J(\mu, V, H) - J(\mu, V, L)] \\ = \; & J(\mu, V, L) + \mu [J(\mu, V, H) - J(\mu, V, L)] = J(\mu, V), \end{aligned}$$

and $J(\mu, V)$ is convex.

Part 3. Monotonicity with respect to μ.

Suppose that $J(\mu, V) = J$ for some μ and V. Then, for a higher value $\mu' > \mu$ and the same promised utility V, we have $J(\mu', V) \geq J$. To see why, if $(p_z, \psi_z, e_z, V'_z)_z$ satisfies the promise keeping and incentive compatibility for μ, then it satisfies them for μ' as well since none of these constraints depends on μ.

Recall that the instantaneous utility for P given ψ_z and e_z is

$$u^P(\mu, \psi_z, e_z) = -\mu \mathbb{E}[C|\psi_z, e_z] - (1-\mu) \mathbb{E}[C|\psi_z, e_L = 0].$$

Since the expected cost satisfies $\mathbb{E}[C|\psi_z, e_z] \leq \mathbb{E}[C|\psi_z, e_L = 0]$, the instantaneous utility is (weakly) increasing in μ. Hence, by the standard arguments –see Stokey (1989)– $J(\mu, V)$ is (weakly) increasing in μ.

To show it is strictly increasing for $V \in (0,1)$, it suffices to show that $e_z > 0$ for some z if $V \in (0,1)$. The following claim is sufficient: Given the public history h^t, there exists z_t such that the regulator exerts positive effort unless the regulator is replaced at the beginning of period t after each z_t (this corresponds to $V = 0$) or the regulator is kept forever (this corresponds to $V = 1$).

First, note that there must exist $\tilde{t} \geq t$ and $z_{\tilde{t}}$ such that $h^{\tilde{t}}$ happens with a positive probability, the same regulator stays until period $\tilde{t}$ given $h^{\tilde{t}}$, and $e_{z_{\tilde{t}}} > 0$. Otherwise, the social welfare would be $\underline{SW}$; however, since $V \in (0,1)$, providing a small reward in terms of continuation payoff after a good outcome is feasible. Since $c_H'(0) = 0$, a small reward is enough to implement a positive effort and obtain welfare higher than $\underline{SW}$, which is a contradiction.

Second, if $\tilde{t} > t$, then from period t on, rather than playing $\sigma_P|h^t$ and $\sigma_R|h^t$ (the equilibrium strategy given history h^t), suppose P and R play $\sigma_P|h^{\tilde{t}}$ and $\sigma_R|h^{\tilde{t}}$, as if the current period were period $\tilde{t}$ and the history were $h^{\tilde{t}}$. If $\sigma_P|h^t$ and $\sigma_R|h^t$ is a sequential equilibrium, then so is $\sigma_P|h^{\tilde{t}}$ and $\sigma_R|h^{\tilde{t}}$. By discounting, front-loading the effort improves P's welfare. In the sustainable equilibrium (best sequential equilibrium for the public), this is a contradiction. Therefore, $\tilde{t} = t$, and there exists z_t such that R exerts positive effort, as desired.

A.1.6 Proof of Lemma 3

We have $J(\mu, 0) = \bar{J}$ for each μ since the public has to replace R right away. Hence we are left to prove the other two properties:

Part 1. There exists $V(\mu)$ such that $J(\mu, V)$ is linear for $V \in [0, V(\mu)]$.

Suppose such $V(\mu)$ does not exist. By Lemma 2, this means that $J(\mu, V)$ is strictly concave near $V = 0$.

Take $V \in (0, 1-\delta)$. This means that the public needs to stochastically replace R, since otherwise R receives $1-\delta$ by not working. Let β be the probability of a replacement. The promise keeping condition implies

$$\beta \times 0 + (1-\beta) \times \hat{V} = V, \tag{22}$$

where $\hat{V} \geq 1-\delta$ is the promised utility conditional on R not being replaced.

P maximizes

$$\max_{\beta, \hat{V}} \beta J(\mu, 0) + (1-\beta) J\left(\mu, \hat{V}\right)$$

subject to

$$\beta \times 0 + (1-\beta) \times \hat{V} = V \text{ and } \hat{V} \geq 1-\delta.$$

Substituting the constraint, P's welfare is

$$J(\mu, 0) + \frac{V}{\hat{V}}\left[J\left(\mu, \hat{V}\right) - J(\mu, 0)\right]. \tag{23}$$

Taking the derivative with respect to $\hat{V}$ (the differentiability of $J\left(\mu, \hat{V}\right)$ follows from the Envelope Theorem), we obtain

$$\begin{aligned} & V \frac{J_{\hat{V}}\left(\mu, \hat{V}\right)\hat{V} - \left[J\left(\mu, \hat{V}\right) - J(\mu, 0)\right]}{\left(\hat{V}\right)^2} \\ = \; & V \frac{J(\mu, 0) + \left[J_{\hat{V}}\left(\mu, \hat{V}\right)\hat{V} - J\left(\mu, \hat{V}\right)\right]}{\left(\hat{V}\right)^2}. \end{aligned}$$

At $\hat{V} = 0$, this expression is zero. Taking the derivative of the numerator, we obtain

$$\begin{aligned} & \frac{d}{d\hat{V}}\left\{J(\mu, 0) + \left[J_{\hat{V}}\left(\mu, \hat{V}\right)\hat{V} - J\left(\mu, \hat{V}\right)\right]\right\} \\ = \; & J_{\hat{V}\hat{V}}\left(\mu, \hat{V}\right)\hat{V}. \end{aligned}$$

Since we assumed $J(\mu, \cdot)$ is strictly concave, this is negative near $\hat{V} = 0$; and since $J(\mu, \cdot)$ is concave, this is always non-positive. Hence, the smallest $\hat{V} = 1-\delta$ is optimal. Given

$\hat{V} = 1 - \delta$, by (23), we have

$$J(\mu, V) = J(\mu, 0) + \frac{V}{1-\delta}\left[J(\mu, 1-\delta) - J(\mu, 0)\right],$$

for $V \in [0, 1-\delta]$, which is linear in V.

Part 2. For $\mu \geq \mu_H$, we have $V(\mu) > 1 - \delta$.

Suppose $\mu \geq \mu_H$. For the sake of contradiction, assume that $V \leq 1 - \delta$ for each $V \in \arg\max_V J(\mu, V)$. Then, in the above problem, $\hat{V} = 1 - \delta$ is the unique optimum. Recall that β is defined as the probability of a replacement in (22). Hence P cannot replace R in the current period after P picks $\hat{V}$ with probability $1 - \beta$. If P promised a positive continuation payoff from the next period, then R could obtain a payoff greater than $1 - \delta$ without working at all. We therefore have to make sure that $V'[\hat{V}](o_z) = 0$ for each z and o_z, and so $e_z = 0$. Therefore, the effort has to be equal to 0. Then, the public's instantaneous welfare is $(1-\delta)\underline{SW}$. Moreover, since $V'[\hat{V}](o) = 0$ for each o, R will be replaced in the next period with probability one. Hence, the continuation social welfare is $\delta\bar{J}$. In total,

$$J(\mu, 1-\delta) = (1-\delta)\underline{SW} + \delta\bar{J}. \tag{24}$$

For $\mu = \mu_H$, (24) together with (21) implies that $J(\mu, 0) = \bar{J}$ and $J(\mu, V)$ is linear and less than $\bar{J}$ for each $V \in (0, 1-\delta]$. By concavity, this means that $J(\mu, V) < \bar{J}$ for each $V > 0$. Thus, $\arg\max_V J(\mu, V) = 0$. This means that $\bar{J}$ is uniquely obtained by always replacing R; however, this implies that R exerts no effort, which is a contradiction. Hence, $V(\mu_H) > 1 - \delta$. Moreover, since $\bar{J} = \max_V J(\mu_H, V)$, it follows that

$$J(\mu_H, V) = \bar{J} \text{ for } V \in [0, V(\mu_H)]. \tag{25}$$

For $\mu > \mu_H$, by Lemma 2, we have $J(\mu, 1-\delta) > J(\mu_H, 1-\delta) \geq \bar{J}$, which contradicts (24). Hence, $V(\mu) > 1 - \delta$ as well.

Part 3. The Slope of the Linear Part.

Since $J(\mu, V)$ is strictly increasing in $\mu \in (0,1)$, and $J(\mu, 0) = \bar{J}$ for each μ, (25) implies the slope of the linear part is negative for $\mu < \mu_H$ and positive for $\mu > \mu_H$.

Part 4. Property of $V \in \arg\max_{\hat{V}} J\left(\mu, \hat{V}\right)$

Without loss of generality, we can take $V \in \arg\max_{\hat{V}} J\left(\mu, \hat{V}\right)$ such that V is the extreme point of the graph $\left\{\hat{V}, J\left(\mu, \hat{V}\right)\right\}_{\hat{V}}$. This means that no mixture can implement V. Hence, the social welfare $J(\mu, V)$ at $V \in \arg\max_{\hat{V}} J\left(\mu, \hat{V}\right)$, denoted by $J(\mu)$, is determined by the dynamic program without mixture:

$$\begin{aligned} J(\mu) &= \max_{(\psi, e, V')} \{(1-\delta) u^P(\mu, \psi, e) \\ &\quad + \delta \sum_o \Pr(o|\mu, \psi, e) J(\mu'(\mu, \psi, e, o), V'(o))\}, \end{aligned}$$

subject to incentive compatibility constraint:

$$e \in \arg\max \left\{(1-\delta)[1 - c_H(e)] + \delta \sum_o \Pr(o|\psi, e) V'(o)\right\}.$$

Note that we do not impose the promise keeping constraint since we are free to choose $\hat{V}$ to maximize $J\left(\mu, \hat{V}\right)$. Moreover, since the first-order condition for e is always necessary and sufficient by the assumption of the cost function c, we can see the above dynamic programming as deciding $(V'(o))_o$, and then e is determined by the first-order condition.

In this problem, we first show that $V'(o) \leq \arg\max_{\hat{V}} J\left(\mu'(\mu, \psi, e, o), \hat{V}\right)$ after $\mu'(\mu, \psi, e, o) \leq \mu$. Suppose otherwise: There exists $\bar{o}$ such that $V'(\bar{o}) > \arg\max_{\hat{V}} J\left(\mu'(\mu, \psi, e, \bar{o}), \hat{V}\right)$ after $\mu'(\mu, \psi, e, \bar{o}) \leq \mu$.

Since

$$\begin{aligned} \mu'(\mu, \psi, e, \bar{o}) &= \frac{\mu \Pr(\bar{o}|\psi, e)}{\mu \Pr(\bar{o}|\psi, e) + (1-\mu) \Pr(\bar{o}|\psi, e=0)} \\ &= \frac{\mu}{\mu + (1-\mu) \frac{\Pr(\bar{o}|\psi, e=0)}{\Pr(\bar{o}|\psi, e)}} \leq \mu, \end{aligned} \tag{26}$$

we have $\Pr(\bar{o}|\psi, 0) \geq \Pr(\bar{o}|\psi, e)$. $\Pr(o|\psi, e)$ is monotone in e for each o and ψ, so the probability $\Pr(\bar{o}|\psi, e)$ is decreasing in e.

Then, the first-order condition for the optimality of $V'(\bar{o})$ is

$$\begin{aligned} 0 &= \frac{d}{dV'(\bar{o})}\{(1-\delta)u^P(\mu,\psi,e) \\ &\quad +\delta\sum_o \Pr(o|\mu,\psi,e) J(\mu'(\mu,\psi,e,o), V'(o))\} \\ &= \{(1-\delta)u_e^P(\mu,\psi,e) \\ &\quad +\delta\sum_o [\frac{d}{de}\Pr(o|\mu,\psi,e)] J(\mu'(\mu,\psi,e,o), V'(o)) \\ &\quad +\delta\sum_o \Pr(o|\mu,\psi,e) J_1(\mu'(\mu,\psi,e,o), V'(o))[\frac{d}{de}\mu'(\mu,\psi,e,o)]\}\frac{de}{dV'(\bar{o})} \\ &\quad +\delta \Pr(\bar{o}|\mu,\psi,e) J_2(\mu'(\mu,\psi,e,\bar{o}), V'(\bar{o})), \end{aligned}$$

where J_n is the derivative of J with respect to its nth argument; and $u_e^P \geq 0$ is the derivative of u^P with respect to e. Since $\Pr(\bar{o}|\psi, e)$ is decreasing in e, it follows that $\frac{de}{dV'(\bar{o})} < 0$. Moreover, $J_2(\mu'(\mu,\psi,e,\bar{o}), V'(\bar{o})) < 0$, since $V'(\bar{o}) > \arg\max_{\hat{V}} J\left(\mu'(\mu,\psi,e,\bar{o}), \hat{V}\right)$ and J is concave. Hence, we have

$$\begin{aligned} &\{(1-\delta)u_e^P(\mu,\psi,e) + \delta\sum_o [\frac{d}{de}\Pr(o|\mu,\psi,e)] J(\mu'(\mu,\psi,e,o), V'(o)) \\ &+\delta\sum_o \Pr(o|\mu,\psi,e) J_1(\mu'(\mu,\psi,e,o), V'(o))[\frac{d}{de}\mu'(\mu,\psi,e,o)]\} < 0. \end{aligned} \tag{27}$$

Similarly, if there exists $\hat{o}$ such that $\Pr(\hat{o}|\psi, e)$ is decreasing in e but

$$V'(\hat{o}) \leq \arg\max_{\hat{V}} J\left(\mu'(\mu,\psi,e,\hat{o}), \hat{V}\right),$$

then the symmetric argument implies that

$$\begin{aligned} &\{(1-\delta)u_e^P(\mu,\psi,e) + \delta\sum_o \left[\frac{d}{de}\Pr(o|\mu,\psi,e)\right] J(\mu'(\mu,\psi,e,o), V'(o)) \\ &+\delta\sum_o \Pr(o|\mu,\psi,e) J_1(\mu'(\mu,\psi,e,o), V'(o))[\frac{d}{de}\mu'(\mu,\psi,e,o)]\} \geq 0, \end{aligned}$$

which is a contradiction.

Therefore, letting O_- be the set of outcomes o such that $\Pr(o|\psi, e)$ is decreasing in e, for each $o \in O_-$, we have $V'(o) > \arg\max_{\hat{V}} J\left(\mu'(\mu, \psi, e, o), \hat{V}\right)$. Symmetrically, letting O_+ be the set of o such that $\Pr(o|\psi, e)$ is increasing in e, for each $o \in O_+$, we have $V'(o) < \arg\max_{\hat{V}} J\left(\mu'(\mu, \psi, e, o), \hat{V}\right)$.

Now we set $V^*(o) = \arg\max_{\hat{V}} J\left(\mu'(\mu, \psi, e, o), \hat{V}\right)$ for each o, and let e^* be the new optimal effort (fixing ψ throughout). Since $V^*(o) < V'(o)$ for $o \in O_-$ and $V^*(o) > V'(o)$ for $o \in O_+$, we have $e^* > e$ (here, e is the original effort). Hence, we have

$$(1-\delta)\, u^P(\mu, \psi, e^*) > (1-\delta)\, u^P(\mu, \psi, e). \tag{28}$$

In addition, we adjust $V^*(o)$ so that the continuation welfare increases with fixed e:

$$\begin{aligned} & \sum_o \Pr(o|\mu, \psi, e)\, J\left(\mu'(\mu, \psi, e, o), V'(o)\right) \\ < & \sum_o \Pr(o|\mu, \psi, e)\, J\left(\mu'(\mu, \psi, e, o), V^*(o)\right). \end{aligned} \tag{29}$$

Moreover, since $\max_{\hat{V}} J\left(\mu', \hat{V}\right)$ is increasing in μ', we have

$$J\left(\mu'(\mu, \psi, e, o), V^*(o)\right) < J\left(\mu'(\mu, \psi, e, \hat{o}), V^*(\hat{o})\right)$$

for each $o \in O_-$ and $\hat{o} \in O_+$. Since e^* increases the probability of event o if and only if $o \in O_+$, we have

$$\begin{aligned} & \sum_o \Pr(o|\mu, \psi, e)\, J\left(\mu'(\mu, \psi, e, o), V^*(o)\right) \\ < & \sum_o \Pr(o|\mu, \psi, e^*)\, J\left(\mu'(\mu, \psi, e, o), V^*(o)\right). \end{aligned} \tag{30}$$

Finally, learning (the difference between $\mu'(\mu, \psi, e, o)$ and $\mu'(\mu, \psi, e^*, o)$) further increases the continuation payoff. To show this, we first make the following claim:

Claim 1 *For* $\mu_1 < \mu_2$, $V^*(\mu_1) \in \arg\max_{\hat{V}} J\left(\mu_1, \hat{V}\right)$ *and* $V^*(\mu_2) \in \arg\max_{\hat{V}} J\left(\mu_2, \hat{V}\right)$, *the following holds:*

$$J_1(\mu_1, V^*(\mu_1)) \leq J_1(\mu_2, V^*(\mu_2)).$$

Proof. We have

$$\begin{aligned} & J(\mu_1, V^*(\mu_1)) + J_1(\mu_1, V^*(\mu_1))[\mu_2 - \mu_1] \\ \leq\ & J(\mu_2, V^*(\mu_1)) \text{ since } J \text{ is convex in } \mu \\ \leq\ & J(\mu_2, V^*(\mu_2)) \text{ since } V^*(\mu_2) \text{ is } \arg\max \text{ for } \mu_2. \end{aligned}$$

At the same time,

$$\begin{aligned} J(\mu_1, V^*(\mu_1)) &\geq J(\mu_2, V^*(\mu_2)) - J_1(\mu_2, V^*(\mu_2))[\mu_2 - \mu_1] \text{ since } J \text{ is convex in } \mu \\ &\geq J(\mu_1, V^*(\mu_1)) + J_1(\mu_1, V^*(\mu_1))[\mu_2 - \mu_1] \\ &\quad - J_1(\mu_2, V^*(\mu_2))[\mu_2 - \mu_1] \text{ from the previous inequality.} \end{aligned}$$

Hence,

$$0 \geq [J_1(\mu_1, V^*(\mu_1)) - J_1(\mu_2, V^*(\mu_2))](\mu_2 - \mu_1).$$

■

Given this claim, since $J_1(\mu'(\mu, \psi, e, o), V^*(o))$ is larger for o with $\mu'(\mu, \psi, e, o) > \mu$ than for o with $\mu'(\mu, \psi, e, o) < \mu$, increasing (decreasing) $\mu'(\mu, \psi, e^*, o) > \mu'(\mu, \psi, e, o)$ for o with $\mu'(\mu, \psi, e, o) > \mu$ (for o with $\mu'(\mu, \psi, e, o) < \mu$) increases the continuation payoff since the distribution of $\{\mu'(\mu, \psi, e^*, o)\}_o$ given e^* is the mean-preserving spread of that of $\{\mu'(\mu, \psi, e, o)\}_o$ given e. Together with (29) and (30), this leads to

$$\begin{aligned} & \sum_o \Pr(o|\mu, \psi, e) J(\mu'(\mu, \psi, e, o), V'(o)) \\ <\ & \sum_o \Pr(o|\mu, \psi, e^*) J(\mu'(\mu, \psi, e^*, o), V^*(o)). \end{aligned}$$

Together with (28), we have proven that the social welfare increases.

The proof for $V'(o) \geq \arg\max_{\hat{V}} J\left(\mu'(\mu,\psi,e,o), \hat{V}\right)$ after $\mu'(\mu,\psi,e,o) \geq \mu$ is completely symmetric, and so it is omitted.

A.1.7 Proof of Lemmas 4 and 5

Part 1. With separated institutions:

From (16), R's choice of effort satisfies

$$(1-\delta)\,c'_H(e_z) = \delta q'(e_z)\left[V'_z(0|\psi=0) - V'_z(C_2|\psi=0)\right].$$

It then follows that

$$V'_z(0|\psi=0) - V'_z(C_2|\psi=0) = \frac{1-\delta}{\delta}\frac{c'_H(e_z)}{q'(e_z)} > 0. \tag{31}$$

Given O's strategy,(15) becomes:

$$V_z = (1-\delta)\left[u - c(e_z)\right] + \delta q(e_z)\left[V'_z(0|\psi=0) - V'_z(C_2|\psi=0)\right] + \delta V'_z(C_2|\psi=0).$$

Then, given (31), $V'_z(C_2|\psi=0)$ is a function of both e_z and V_z:

$$\delta V'_z(C_2|\psi=0) = V_z - (1-\delta)\left[u - c(e_z) + q(e_z)\frac{c'_H(e_z)}{q'(e_z)}\right],$$

and so

$$\frac{\partial V'_z(C_2|\psi=0)}{\partial e_z} < 0. \tag{32}$$

Since (31) implies

$$\frac{\partial\left[V'_z(0|\psi=0) - V'_z(C_2|\psi=0)\right]}{\partial e_z} > 0,$$

we have

$$\frac{\partial V'_z(0|\psi=0)}{\partial e_z} > 0. \tag{33}$$

Part 2. With the unified institution:

From (16),

$$(1-\delta)\, c_H'(e_z) = \delta(1-\varepsilon)\, q'(e_z)\left[V_z'(0) - V_z'(C_1)\right],$$

so

$$V_z'(0|\psi=1) - V_z'(C_1|\psi=1) = \frac{1-\delta}{\delta}\frac{c_H'(e_z)}{(1-\varepsilon)\, q'(e_z)}. \tag{34}$$

Since O intervenes if and only if $s=B$, the promise keeping (15) becomes

$$V_z = (1-\delta)\left[u - c(e_z)\right] + \delta(1-\varepsilon)\, q(e_z)\left[V_z'(0) - V_z'(C_1)\right] + \delta V_z'(C_1),$$

so

$$\delta V_z'(C_1|\psi=1) = V_z - (1-\delta)\left[u - c(e_z) + q(e_z)\frac{c_H'(e_z)}{q'(e_z)}\right].$$

As in the case with separated institutions, we have

$$\frac{\partial V_z'(C_1|\psi=1)}{\partial e_z} < 0 \text{ and } \frac{\partial V_z'(0|\psi=1)}{\partial e_z} > 0. \tag{35}$$

Then, (32), (33), and (35) establish Lemma 4, and (31) and (34) establish Lemma 5.

A.1.8 Proof of Proposition 4 and Corollary 2

Consider the difference in belief updates after the worst outcome for each institution:

$$\Delta(\mu) = \frac{\mu\left[1-q(e)\right]}{1-\left[\mu q(e) + (1-\mu)\, q(0)\right]} - \frac{\mu\left[1-(1-\varepsilon)\, q(e)\right]}{1-(1-\varepsilon)\left[\mu q(e) + (1-\mu)\, q(0)\right]}.$$

The similar calculation shows that the belief updates after the good outcome $C=0$ is the same regardless of the institution structure.

Since $\Delta(\mu) \geq 0$ for each μ, Proposition 4 holds. Taking the second derivative with respect to μ:

$$\frac{\partial^2 \Delta(\mu)}{\partial \mu^2} < 0,$$

showing that Δ is a concave function of μ. At $\mu = 0$, we have $\Delta = 0$; and at $\mu = 1$, we have $\Delta = 0$. Hence, we have $\Delta > 0 \ \forall \mu \in (0, 1)$. This also implies that $\exists \mu^* \in (0, 1)$ such that $\mu^* = \arg\max \Delta(\mu)$.

A.2 Additional Results

A.2.1 Generalization of the Regulator's Utility

One may wonder whether the static trade-off obtained in Section 3 is due to the form of the regulator's utility function, which limits the benefits of reputation to only one outcome. In this section, we show that the model is in fact robust to more general utility forms, in which R's reputation changes continuously as a function of the possible outcomes.

Consider the general form for R's utility, $u^R(b|C)$, where $u^R(\cdot) \geq 0$ is a concave increasing function of $b|C \equiv \Pr(\theta = H|C)$. We still have an equilibrium in which the regulator of each type exerts zero effort. Again, we focus on the equilibrium in which type H exerts positive effort.

Separated Institutional Structure The general form of $u^R(\cdot)$ does not significantly alter the properties of the equilibrium.

Proposition 6 *There exists an equilibrium with positive effort and separate institutions. The equilibrium satisfies the following properties:*

- ***(Effort strategy)*** *The equilibrium effort choices satisfies $e_H > e_L \geq 0$.*
- ***(Intervention Strategy)*** *O implements $\iota = 0$.*
- ***(Beliefs)*** *The equilibrium beliefs satisfy:*

$$\Pr(\theta = H|C = 0) > \mu > \Pr(\theta = H|C = C_2).$$

Proof. The same as Proposition 1 (note that the proof in Section A.1.1 does not depend on the form of the utility function). ■

As before, type H exerts higher effort in equilibrium. Given Assumption 1, O does not intervene without additional information. Finally, as before, the public's beliefs about R's type are highest after a good outcome ($C = 0$) and lowest after a crisis ($C = C_2$). The only difference from Proposition 1 is that the equilibrium with positive effort is no longer unique. Since the return on effort now continuously depends on reputation and reputation in turn depends on the (rational) expectation of the equilibrium effort, the structure of the game is one of a "coordination-game," which may give rise to multiple equilibria.

With the general form of the utility function, the first-order condition (10) becomes

$$q'(e_\theta)\left(u^R(b|C=0) - u^R(b|C=1)\right) = c'_\theta(e_\theta). \tag{36}$$

In turn, $b|C$ depends on the equilibrium effort $\{e_\theta\}_\theta$. For each e_H, let $e_L(e_H)$ be the low type's effort such that

$$\frac{q'(e_H)}{q'(e_L)} = \frac{c'_H(e_H)}{c'_L(e_L)}. \tag{37}$$

Since q is concave and c_θ is convex, there is a unique solution. Intuitively, if e_H is the equilibrium effort of type H, then (36) holds for both types if and only if type L's effort is $e_L(e_H)$. Given e_H and $e_L(e_H)$, let $b_C(e_H)$ be the belief:

$$\begin{aligned} b_0(e_H) &= \frac{\mu q(e_H)}{\mu q(e_H) + (1-\mu) q(e_L(e_H))}; \\ b_{C_2}(e_H) &= \frac{\mu[1-q(e_H)]}{\mu[1-q(e_H)] + (1-\mu)[1-q(e_L(e_H))]}. \end{aligned}$$

Then, the equilibrium condition boils down to

$$q'(e_H)\left(u^R(b_0(e_H)) - u^R(b_{C_2}(e_H))\right) = c'_H(e_H). \tag{38}$$

Unified Institutional Structure In the unified institution, again, the equilibrium properties described in Proposition 2 are maintained.

Proposition 7 *There exists an equilibrium with positive effort and one unified institution. The equilibrium satisfies the following properties:*

- ***(Effort strategy)*** *The equilibrium effort choices satisfied $e_H > e_L \geq 0$.*

- ***(Intervention Strategy)*** *O intervenes after $s = B$ and does not intervene after $s = G$.*

- ***(Beliefs)*** *The equilibrium beliefs satisfy:*

$$\Pr(\theta = H|C = 0) > \mu > \Pr(\theta = H|C = C_1) \geq \Pr(\theta = H|C = C_2).$$

Proof. The same as Proposition 2. ■

As before, R's effort is strictly higher for type H whenever positive effort is undertaken. With the general form of the utility function, the first-order condition (11) becomes

$$q'\left(e_\theta^\varepsilon\right)\left\{(1-\varepsilon)\left(u(b^\varepsilon|0) - u\left(b^\varepsilon|C_2\right)\right) + \varepsilon\left(u\left(b^\varepsilon|C_1\right) - u(b^\varepsilon|C_2)\right)\right\} = c'_\theta\left(e_\theta^\varepsilon\right).$$

Superscript ε denotes the probability of error. As above, if e_H^ε is the equilibrium effort of type H, then type L's effort must be $e_L\left(e_H^\varepsilon\right)$. Given e_H and $e_L\left(e_H\right)$, let $b_C^\varepsilon\left(e_H\right)$ be the belief:

$$\begin{aligned}
b_0^\varepsilon\left(e_H\right) &= \frac{\mu q\left(e_H\right)}{\mu q\left(e_H\right) + (1-\mu)\, q\left(e_L\left(e_H\right)\right)}; \\
b_{C_1}^\varepsilon\left(e_H\right) &= \frac{\varepsilon\mu q\left(e_H\right) + (1-\varepsilon)\,\mu\left[1 - q\left(e_H\right)\right]}{\left\{\begin{array}{c} \varepsilon\mu q\left(e_H\right) + (1-\varepsilon)\,\mu\left[1 - q\left(e_H\right)\right] \\ +\varepsilon(1-\mu)\, q\left(e_L\left(e_H\right)\right) + (1-\varepsilon)(1-\mu)\left[1 - q\left(e_L\left(e_H\right)\right)\right] \end{array}\right\}}; \\
b_{C_2}^\varepsilon\left(e_H\right) &= \frac{\mu\left[1 - q\left(e_H\right)\right]}{\mu\left[1 - q\left(e_H\right)\right] + (1-\mu)\left[1 - q\left(e_L\left(e_H\right)\right)\right]}.
\end{aligned} \tag{39}$$

The equilibrium condition is

$$q'\left(e_H^\varepsilon\right)\left\{\begin{array}{c} (1-\varepsilon)\left(u(b_0^\varepsilon\left(e_H^\varepsilon\right)) - u\left(b_{C_2}^\varepsilon\left(e_H^\varepsilon\right)\right)\right) \\ +\varepsilon\left(u\left(b_{C_1}^\varepsilon\left(e_H^\varepsilon\right)\right) - u\left(b_{C_2}^\varepsilon\left(e_H^\varepsilon\right)\right)\right) \end{array}\right\} = c'_H\left(e_H^\varepsilon\right). \tag{40}$$

Ex-ante and Ex-post Trade-offs. Since (36) in the separated institutional structure corresponds to (40) with $\varepsilon = 0$, condition (40) can be used to derive the following comparative static of effort with respect to ε. Suppose $\varepsilon > 0$ and consider an equilibrium satisfying

(40). Suppose we increase ε (moving away from the separated institution). Since $b_0^\varepsilon(e_H^\varepsilon) > b_{C_1}^\varepsilon(e_H^\varepsilon) > b_{C_2}^\varepsilon(e_H^\varepsilon)$ by Lemma 7, keeping the public's belief fixed, the increase in ε reduces the reward for regulatory effort,

$$(1-\varepsilon)\left(u(b_0^\varepsilon(e_H^\varepsilon)) - u\left(b_{C_2}^\varepsilon(e_H^\varepsilon)\right)\right) + \varepsilon\left(u\left(b_{C_1}^\varepsilon(e_H^\varepsilon)\right) - u\left(b_{C_2}^\varepsilon(e_H^\varepsilon)\right)\right). \tag{41}$$

This decreases both types' effort provision. Let $\bar{e}_\theta^\varepsilon$ be the new effort, fixing the public's belief.

As a result, the belief $b_C^\varepsilon(e_H^\varepsilon)$ now changes. If this change in the belief decreases the reward (41), then this further decreases the incentive to exert effort, and we can conclude, by the monotone comparative statics, that the equilibrium effort is lower with larger ε. On the other hand, if (41) increases, then it may increase the incentive to exert effort. If this effect is sufficiently strong to overturn the initial decrease in effort, then effort may increase.

To ensure that effort decreases for sure, a sufficient condition is that the public's beliefs does not change drastically as a result of the effort changing to $\bar{e}_\theta^\varepsilon$ (and so the subsequent effect does not overturn the initial decrease in effort): compared to the original belief, with $\bar{e}_\theta^\varepsilon$, the public does not put much higher probability on R being type L if the outcome $C = 0$ is not observed. If this were not the case, the reputational gains from more effort–or the reputational losses from an intervention or a bad outcome–would lead R to want to put in ever more effort to avoid the risk of his reputation decreasing.

Mathematically, it is sufficient to assume that $c'_H(e_H)$ is convex and

$$MR_\varepsilon(e_H) \equiv q'(e_H)\left\{\begin{array}{c}(1-\varepsilon)\left(u^R(b_0^\varepsilon(e_H)) - u^R\left(b_{C_1}^\varepsilon(e_H)\right)\right) \\ -\varepsilon\left(u^R(b_{C_2}^\varepsilon(e_H)) - u^R\left(b_{C_1}^\varepsilon(e_H)\right)\right)\end{array}\right\} \tag{42}$$

is concave in e_H.

Given the convexity of $c'_H(e_H)$ and concavity of $MR_\varepsilon(e_H)$, we can now show that the same trade-offs highlighted in the simple model from Section 3 exist in the model with the general payoff from reputation:

Proposition 8 *Suppose that $c'_H(e_H)$ is convex and $MR_\varepsilon(e_H)$ is concave in e_H. Then the* *following properties hold:*

1. *In both institutional structures, there exists a unique equilibrium with a positive effort by type H.*

2. *Effort e_H is higher with separate institutions.*

3. *The expected cost is lower under separate institutions if ε and $\frac{C_1}{C_2}$ are sufficiently high.*

Moreover, there exist $q(e)$ and $\{c_\theta(e)\}_\theta$ such that $c'_H(e_H)$ is convex and $MR_\varepsilon(e_H)$ is concave in e_H.

Proof. In Section A.3.1. ■

A.2.2 Integrated Objectives

So far, we have examined the case in which the objective of R and that of O do not change with the institutional structure. One may argue that institutional unification should bring about a change in the objective of R or O. In what follows, we study this case.

We consider the same setup as in Section 3, with the change that now the utility of R in the unified institution is given by:

$$u_\theta^{uni} = \alpha\left[u^R(\Pr(\theta = H|C)) - c_\theta(e)\right] + (1-\alpha)(-C),$$

where $\alpha \in (0,1)$. That is, in the unified institution, R also cares about O. On the other hand, in the separate institution, R's utility is as before:

$$u_\theta^{sep} = u^R(\Pr(\theta = H|C)) - c_\theta(e).$$

The objective of O remains that of minimizing the expected cost:

$$u = -C.$$

Again, given Assumption 1, O intervenes after $s = B$ and does not intervene after $s = G$ in the unified institution; and O does not intervene in the separated institution. Here, we keep O's objective unchanged. A similar analysis can show the robustness of our result to the case in which O cares about R.

The problem and the equilibrium characterization does not change in the case of separated institutions. In the case of institutional unification, R maximizes his objective

$$\max_{e} \alpha \left[\Pr(y = G, s = G|e) - c_{\theta}(e) \right] + \\ (1 - \alpha) \left(- \Pr(s = B|e) C_1 - \Pr(s = G, y = B|e) C_2 \right),$$

which leads to the first-order condition

$$q'(e) \left[(1 - \varepsilon) + \frac{1 - \alpha}{\alpha} \left(\varepsilon C_2 - \varepsilon C_1 + (1 - \varepsilon) C_1 \right) \right] = c'_{\theta}(e).$$

Analyzing this condition, it follows that R's effort choice is higher whenever he places some positive weight on O's objective. Intuitively, the higher effort always benefits O by reducing the probability of bad outcomes (either C_1 or C_2). The trade-offs described in Section 3 persist, however, in the model with integrated regulatory objective, as long as R continues to place sufficient weight on reputation.

Result 1 *In the unified institution, regulatory effort decreases in α. Moreover, there exists $\alpha^* \in (0, 1)$ such that, for each $\alpha > \alpha^*$,*

1. *R's effort is higher in the separated institution.*

2. *The expected cost is larger with the unified institution than with separate institutions if ε and $\frac{C_1}{C_2}$ are sufficiently high.*

Proof. In Section A.3.2. ■

A.3 Dynamic analysis when type L also works

Lemma 1 is readily extended to the case in which type L works. Hence, we are left to characterize the on-path strategy by dynamic programming.

Once P decides the replacement probability after each event, it determines the best response of type L. In particular, once P decides the target effort for type H, e_z^H, and the continuation promised utility for type H, $V'(o_z)$, type L maximizes his own payoff:

$$\begin{aligned} V_L(\mu, V) &= \max_{\{e_z\}_z} \int_z (1-p_z)\{(1-\delta)(1-c_L(e_z)) \\ &\quad +\delta \sum_{o_z} \Pr(o_z|\psi_z, e_z) V_L\left(\mu'\left(\mu, \psi_z, e_z^L, e_z^H, o_z\right), V'(o_z)\right)\}dz. \end{aligned}$$

Here, $V_L(\mu, V)$ is the value function of the regulator with type L when the belief is μ and the promised utility for type H is V. Now the updated belief μ' depends not only on H's equilibrium effort e_z^H but L's, e_z^L. In the maximization problem, $\left\{e_z^L, e_z^H\right\}_z$ is the equilibrium effort for the two types. Since the public believes that the regulator follows the equilibrium effort when calculating the belief update, effort e_z^L is not controlled by the type L regulator.

The public's problem is to maximize

$$\begin{aligned} J(\mu, V) &= \max_{(p_z, \psi_z, e_z^H, e_z^L, V_z')_z} \int_z [p_z \bar{J} + (1-p_z)\{(1-\delta) u^P\left(\mu, \psi_z, e_z^H, e_z^L\right) \\ &\quad +\delta \sum_{o_z} \Pr(o_z|\mu, \psi_z, e_z) J\left(\mu'\left(\mu, \psi_z, e_z^H, e_z^L, o_z\right), V_z'(o_z)\right)\}]dz, \end{aligned}$$

where

$$u^P\left(\mu, \psi_z, e_z^H, e_z^L\right) = \mu \mathbb{E}\left[-C|\psi_z, e_z^H\right] + (1-\mu)\mathbb{E}\left[-C|\psi_z, e_z^L\right]$$

is P's instantaneous utility. At the same time, we keep track of type L's value function:

$$\begin{aligned} V_L(\mu, V) &= \int_z (1-p_z)\{(1-\delta)\left(1-c_L\left(e_z^L\right)\right) \\ &\quad +\delta \sum_{o_z} \Pr\left(o_z|\psi_z, e_z^L\right) V_L\left(\mu'\left(\mu, \psi_z, e_z^L, e_z^H, o_z\right), V'(o_z)\right)\}dz. \end{aligned}$$

The public is subject to the following constraints:

1. Promise keeping constraint:

$$V = \int_z (1-p_z)\left\{(1-\delta)\left[1-c_H\left(e_z^H\right)\right] + \delta \sum_{o_z} \Pr\left(o_z|\psi_z, e_z^H\right) V_z'(o_z)\right\} dz; \tag{43}$$

2. Incentive compatibility constraint for the high type: For each z with $p_z > 0$,

$$e_z^H \in \arg\max_e \left\{(1-\delta)\left[1-c_H(e)\right] + \delta \sum_{o_z} \Pr(o_z|\psi_z, e)\, V_z'(o_z)\right\}; \tag{44}$$

3. Incentive compatibility constraint for the low type: For each z with $p_z > 0$,

$$e_z^L \in \arg\max_e \left\{ \begin{array}{c} (1-\delta)(1-c_L(e)) \\ +\delta \sum_{o_z} \Pr(o_z|\psi_z, e)\, V_L\left(\mu'\left(\mu, \psi_z, e_z^L, e_z^H, o_z\right), V'(o_z)\right) \end{array} \right\}. \tag{45}$$

Hence, now the problem is recursive on the public's welfare function $J(\mu, V)$ and type L's value function $V_L(\mu, V)$. The rest of the analysis is accordingly generalized.

A.3.1 Proof of Proposition 8

For notational convenience, we define

$$\hat{u}(x) = u\left(\frac{1}{1+x}\right),$$

and

$$\begin{aligned} \hat{b}_0^{\varepsilon}(e_H) &= \frac{\mu q(e_H)}{(1-\mu)\, q(e_L(e_H))}; \\ \hat{b}_0^{\varepsilon}(e_H) &= \frac{\varepsilon \mu q(e_H) + (1-\varepsilon)\,\mu\left[1-q(e_H)\right]}{\varepsilon(1-\mu)\, q(e_L(e_H)) + (1-\varepsilon)(1-\mu)\left[1-q(e_L(e_H))\right]}; \\ \hat{b}_{C_2}^{\varepsilon}(e_H) &= \frac{\mu\left[1-q(e_H)\right]}{(1-\mu)\left[1-q(e_L(e_H))\right]}. \end{aligned} \tag{46}$$

Then, we have $\hat{u}\left(\hat{b}_C^{\varepsilon}(e_H)\right) = u\left(b_C^{\varepsilon}(e_H)\right)$ for each e_H and C. We will work with $\hat{u}\left(\hat{b}_C^{\varepsilon}(e_H)\right)$ instead of $u\left(b_C^{\varepsilon}(e_H)\right)$.

Part 1. Uniqueness

The equilibrium condition (40) is $MR_{\varepsilon}(e_H) = c'_H(e_H)$. If $c'_H(e_H)$ is convex and $MR_{\varepsilon}(e_H)$ is concave, since $c'_H(0) = 0$, there are only two solutions: $e_H = 0$ and $e_H > 0$.

Part 2. e_H is decreasing in ε

From the equilibrium condition $MR_{\varepsilon}(e_H) = c'_H(e_H)$, since we assume $c'_H(e_H)$ is convex and $MR_{\varepsilon}(e_H)$ is concave, it is sufficient to show that $\frac{d}{d\varepsilon}MR_{\varepsilon}(e_H) \leq 0$. We have

$$\begin{aligned}\frac{d}{d\varepsilon}MR_{\varepsilon}(e_H) &= -q'(e_H)\left\{\hat{u}\left(\hat{b}_0^{\varepsilon}(e_H)\right) - \hat{u}\left(\hat{b}_{C_1}^{\varepsilon}(e_H)\right) + \hat{u}\left(\hat{b}_{C_1}^{\varepsilon}(e_H)\right) - \hat{u}\left(\hat{b}_{C_2}^{\varepsilon}(e_H)\right)\right\} \\ &\quad -q'(e_H)(1-2\varepsilon)\hat{u}'\left(\hat{b}_{C_1}^{\varepsilon}(e_H)\right)\frac{d}{d\varepsilon}\hat{b}_{C_1}^{\varepsilon}(e_H)\end{aligned}$$

since $\hat{b}_0^{\varepsilon}(e_H)$ and $\hat{b}_{C_2}^{\varepsilon}(e_H)$ do not depend on ε. The first line of the above expression is negative, so we are left to show that $\frac{d}{d\varepsilon}\hat{b}_{C_1}^{\varepsilon}(e_H) \geq 0$.

Since[22]

$$\begin{aligned}&\frac{d}{d\varepsilon}\frac{\mu\left(q(e_H)\varepsilon + (1-q(e_H))(1-\varepsilon)\right)}{(1-\mu)\left(q(e_L)\varepsilon + (1-q(e_L))(1-\varepsilon)\right)} \\ =\ & \frac{\mu}{1-\mu}\frac{\left\{\begin{array}{c}[q(e_H)-(1-q(e_H))][q(e_L)\varepsilon + (1-q(e_L))(1-\varepsilon)] \\ -[q(e_H)\varepsilon + (1-q(e_H))(1-\varepsilon)][q(e_L)-(1-q(e_L))]\end{array}\right\}}{[q(e_L)\varepsilon + (1-q(e_L))(1-\varepsilon)]^2} \\ =\ & \frac{q(e_H)-(1-q(e_H))}{q(e_L)\varepsilon + (1-q(e_L))(1-\varepsilon)}\frac{\mu}{1-\mu} - \frac{q(e_L)-(1-q(e_L))}{q(e_L)\varepsilon + (1-q(e_L))(1-\varepsilon)}\hat{b}_{C_1}^{\varepsilon}\end{aligned}$$

and

$$\frac{q(e_L)-(1-q(e_L))}{q(e_L)\varepsilon + (1-q(e_L))(1-\varepsilon)} \leq \frac{q(e_H)-(1-q(e_H))}{q(e_L)\varepsilon + (1-q(e_L))(1-\varepsilon)},$$

[22] In the algebra below, we simply write e_L instead of $e_L(e_H)$.

we have

$$\begin{aligned}
&\frac{d}{d\varepsilon}\frac{\mu\left(q\left(e_H\right)\varepsilon+\left(1-q\left(e_H\right)\right)\left(1-\varepsilon\right)\right)}{\left(q\left(e_L\right)\varepsilon+\left(1-q\left(e_L\right)\right)\left(1-\varepsilon\right)\right)}\\
\geq\ & \frac{q\left(e_H\right)-\left(1-q\left(e_H\right)\right)}{q\left(e_L\right)\varepsilon+\left(1-q\left(e_L\right)\right)\left(1-\varepsilon\right)}\left(\frac{\mu}{1-\mu}-\hat{b}^{\varepsilon}_{C_1}\right).
\end{aligned}$$

Since $q\left(e_H\right)\geq q\left(e_L\right)$ and $\varepsilon\leq\frac{1}{2}$, we have

$$\begin{aligned}
\frac{\mu}{1-\mu}-\hat{b}^{\varepsilon}_{C_1} &= \frac{\mu}{1-\mu}\left(1-\frac{q\left(e_H\right)\varepsilon+\left(1-q\left(e_H\right)\right)\left(1-\varepsilon\right)}{q\left(e_L\right)\varepsilon+\left(1-q\left(e_L\right)\right)\left(1-\varepsilon\right)}\right)\\
&= \frac{\mu}{1-\mu}\frac{\left(q\left(e_H\right)-q\left(e_L\right)\right)\left(1-2\varepsilon\right)}{q\left(e_L\right)\varepsilon+\left(1-q\left(e_L\right)\right)\left(1-\varepsilon\right)}\geq 0.
\end{aligned}$$

In total, we have

$$\frac{d}{d\varepsilon}\frac{\mu\left(q\left(e_H\right)\varepsilon+\left(1-q\left(e_H\right)\right)\left(1-\varepsilon\right)\right)}{\left(1-\mu\right)\left(q\left(e_L\right)\varepsilon+\left(1-q\left(e_L\right)\right)\left(1-\varepsilon\right)\right)}\geq 0,$$

as desired.

Part 3. Institutional separation is cost-saving if C_2/C_1 is sufficiently small and ε is sufficiently large

Define $c=C_2/C_1$. Let $V\left(\varepsilon,c\right)$ be the welfare with institutional separation minus that with institutional unification, divided by C_1:

$$\begin{aligned}
V\left(\varepsilon,c\right) &= -\sum_{\theta}\mu_\theta\left(1-q\left(e^{sep}_\theta\right)\right)c\\
&\quad+\sum_{\theta}\mu_\theta\underbrace{\left(\left(1-\varepsilon\right)\left(1-q\left(e^{\varepsilon}_\theta\right)\right)+\varepsilon q\left(e^{\varepsilon}_\theta\right)\right)}_{\Pr(s=B|e)}\\
&\quad+\sum_{\theta}\mu_\theta\varepsilon\left(1-q\left(e^{\varepsilon}_\theta\right)\right)c.
\end{aligned}$$

Given c, let $\varepsilon\left(c\right)$ be the value of ε at which $V\left(\varepsilon\left(c\right),c\right)=0$. By the Implicit Function

Theorem,

$$
\begin{aligned}
& -\sum_\theta \mu_\theta \left(1 - q\left(e_\theta^{sep}\right)\right) + \sum_\theta \mu_\theta \begin{pmatrix} -q'\left(e_\theta^{\varepsilon(c)}\right)\left(1 - 2\varepsilon(c)\right)\frac{de_\theta^\varepsilon}{d\varepsilon}\varepsilon'(c) \\ +q\left(e_\theta^{\varepsilon(c)}\right)2\varepsilon'(c) - \varepsilon'(c) \end{pmatrix} \\
& + \sum_\theta \mu_\theta \left(\varepsilon'(c)\left(1 - q\left(e_\theta^{\varepsilon(c)}\right)\right) - \varepsilon(c)\, q'\left(e_\theta^{\varepsilon(c)}\right)\frac{de_\theta^\varepsilon}{d\varepsilon}\varepsilon'(c)\right) c \\
& + \sum_\theta \mu_\theta \varepsilon(c)\left(1 - q\left(e_\theta^{\varepsilon(c)}\right)\right) = 0.
\end{aligned}
$$

$\Leftrightarrow$

$$
\begin{aligned}
& \sum_\theta \mu_\theta \left(\varepsilon(c)\left(1 - q\left(e_\theta^{\varepsilon(c)}\right)\right) - \left(1 - q\left(e_\theta^{sep}\right)\right)\right) \\
= \; & \sum_\theta \mu_\theta \begin{pmatrix} \left(1 - 2\varepsilon(c) + \varepsilon(c)\, c\right) q'\left(e_\theta^{\varepsilon(c)}\right)\frac{de_\theta^\varepsilon}{d\varepsilon} \\ +\left(1 - 2q\left(e_\theta^{\varepsilon(c)}\right)\right) - \left(1 - q\left(e_\theta^{\varepsilon(c)}\right)\right) c \end{pmatrix} \varepsilon'(c).
\end{aligned}
$$

Since $V(\varepsilon(c), c) = 0$,

$$
\begin{aligned}
& \sum_\theta \mu_\theta \varepsilon(c)\left(1 - q\left(e_\theta^{\varepsilon(c)}\right)\right) - \sum_\theta \mu_\theta \left(1 - q\left(e_\theta^{sep}\right)\right) \\
= \; & -\frac{1}{c}\sum_\theta \mu_\theta \left(\left(1 - \varepsilon(c)\right)\left(1 - q\left(e_\theta^{\varepsilon(c)}\right)\right) + \varepsilon(c)\, q\left(e_\theta^{\varepsilon(c)}\right)\right),
\end{aligned}
$$

and so

$$
\sum_\theta \mu_\theta \left(\varepsilon(c)\left(1 - q\left(e_\theta^{\varepsilon(c)}\right)\right) - \left(1 - q\left(e_\theta^{sep}\right)\right)\right) < 0.
$$

Moreover,

$$
\left(1 - 2\varepsilon(c) + \varepsilon(c)\, c\right) q'\left(e_\theta^{\varepsilon(c)}\right)\frac{de_\theta^\varepsilon}{d\varepsilon} \leq 0
$$

from Part 2, and

$$
\begin{aligned}
\left(1 - 2q\left(e_\theta^{\varepsilon(c)}\right)\right) - \left(1 - q\left(e_\theta^{\varepsilon(c)}\right)\right) c \; & \leq \; \left(1 - 2q\left(e_\theta^{\varepsilon(c)}\right)\right) - \left(1 - q\left(e_\theta^{\varepsilon(c)}\right)\right) \\
& \leq \; 0.
\end{aligned}
$$

Therefore,

$$\varepsilon'(c) \geq 0,$$

as desired.

Part 4. When is $MR_\varepsilon(e_H)$ Concave?

A direct calculation gives us

$$\begin{aligned} MR''_\varepsilon(e_H) &= q'''(e_H)\left\{(1-\varepsilon)\left[\hat{u}\left(\hat{b}_0^\varepsilon(e_H)\right)-\hat{u}\left(\hat{b}_{C_1}^\varepsilon(e_H)\right)\right]+\varepsilon\left[\hat{u}\left(\hat{b}_{C_1}^\varepsilon(e_H)\right)-\hat{u}\left(\hat{b}_{C_2}^\varepsilon(e_H)\right)\right]\right\} \\ &\quad +2q''(e_H)\left\{\begin{array}{c}(1-\varepsilon)\left[\hat{u}'\left(\hat{b}_0^\varepsilon(e_H)\right)\hat{b}_0^{\varepsilon\prime}(e_H)-\hat{u}'\left(\hat{b}_{C_1}^\varepsilon(e_H)\right)\hat{b}_{C_1}^{\varepsilon\prime}(e_H)\right] \\ +\varepsilon\left[\hat{u}'\left(\hat{b}_{C_1}^\varepsilon(e_H)\right)\hat{b}_{C_1}^{\varepsilon\prime}(e_H)-\hat{u}'\left(\hat{b}_{C_2}^\varepsilon(e_H)\right)\hat{b}_{C_2}^{\varepsilon\prime}(e_H)\right]\end{array}\right\} \\ &\quad +q'(e_H)\left\{\begin{array}{c}(1-\varepsilon)\left[\begin{array}{c}\hat{u}''\left(\hat{b}_0^\varepsilon(e_H)\right)\left[\hat{b}_0^{\varepsilon\prime}(e_H)\right]^2+\hat{u}'\left(\hat{b}_0^\varepsilon(e_H)\right)\hat{b}_0^{\varepsilon\prime\prime}(e_H) \\ -\hat{u}''\left(\hat{b}_{C_1}^\varepsilon(e_H)\right)\left[\hat{b}_{C_1}^{\varepsilon\prime}(e_H)\right]^2-\hat{u}'\left(\hat{b}_{C_1}^\varepsilon(e_H)\right)\hat{b}_{C_1}^{\varepsilon\prime\prime}(e_H)\end{array}\right] \\ +\varepsilon\left[\begin{array}{c}\hat{u}''\left(\hat{b}_{C_1}^\varepsilon(e_H)\right)\left[\hat{b}_{C_1}^{\varepsilon\prime}(e_H)\right]^2+\hat{u}'\left(\hat{b}_{C_1}^\varepsilon(e_H)\right)\hat{b}_{C_1}^{\varepsilon\prime\prime}(e_H) \\ -\hat{u}''\left(\hat{b}_{C_2}^\varepsilon(e_H)\right)\left[\hat{b}_{C_2}^{\varepsilon\prime}(e_H)\right]^2-\hat{u}'\left(\hat{b}_{C_2}^\varepsilon(e_H)\right)\hat{b}_{C_2}^{\varepsilon\prime\prime}(e_H)\end{array}\right]\end{array}\right\}. \end{aligned}$$

Note that the coefficient of $q'''(e_H)$ is positive. Moreover, the coefficient of $q''(e_H)$ is also positive if $\hat{b}_0^{\varepsilon\prime}(e_H) \geq 0$ (with higher equilibrium effort, the good outcome is better news), $\hat{b}_{C_1}^{\varepsilon\prime}(e_H) \leq 0$ (with higher equilibrium effort, the intervention is worse news), $\hat{b}_{C_1}^{\varepsilon\prime}(e_H) \leq 0$ (with higher equilibrium effort, the bad outcome is worse news), $\hat{b}_{C_2}^{\varepsilon\prime}(e_H) \leq \hat{b}_{C_1}^{\varepsilon\prime}(e_H)$ (the belief after the bad outcome reacts more), $\hat{u}$ is concave.

If these conditions are satisfied and $q'''(e_H) \leq 0$, then it suffices to have the coefficient of $q'(e_H)$ negative. To focus on this issue, let us consider the linear q: $q(e) = q_0 + q_1 e$ (this means that $q''(e) = q'''(e) = 0$, and so if and only if the coefficient of $q'(e_H)$ is negative, then $MR_\varepsilon(e_H)$ is concave). Moreover, for simplicity, assume $\hat{u}$ is linear (note that, if $\hat{u}$ is linear, then $u(b) = -\frac{1}{b}$, which is concave and increasing as desired) and there exists $c(e)$ with $c_\theta(e) = \frac{1}{\theta}c(e)$ for each θ. The latter implies $e_L(e_H) = (L/H)e_H$, by (37). Then the coefficient of $q'(e_H)$ is

$$(1-\varepsilon)\left(\hat{b}_0^{\varepsilon\prime\prime}(e_H)-\hat{b}_{C_1}^{\varepsilon\prime\prime}(e_H)\right)+\varepsilon\left(\hat{b}_{C_1}^{\varepsilon\prime\prime}(e_H)-\hat{b}_{C_2}^{\varepsilon\prime\prime}(e_H)\right). \tag{47}$$

A direct calculation shows that $\hat{b}_C^{\varepsilon}(e_H)$ is concave for each C, and the maximum of (47) is attained with $e_H = 1$. Moreover, the maximum is decreasing in H/L, and for a sufficiently large H/L, (47) becomes negative, as desired.

Intuitively, with a large H/L, if ε goes up and the reward for regulatory effort declines, then type H reacts to the change more than type L. Therefore, the effort supplied by type H gets closer to that supplied by type L. Then, the public does not update its belief much after each event, which further reduces the reward for high effort. Hence, we can conclude that the effort is smaller with institutional unification.

A.3.2 Proof of Result 1

Note that the equilibrium effort level with institutional unification is determined by

$$q'(e)\left[(1-\varepsilon)+\frac{1-\alpha}{\alpha}\left(\varepsilon C_2-\varepsilon C_1+(1-\varepsilon)C_1\right)\right]=c_\theta'(e).$$

On the other hand, the effort level with institutional separation is given by

$$q'(e)=c_\theta'(e).$$

Since q is concave and c_θ is convex, effort is higher with institutional unification if and only if

$$(1-\varepsilon)+\frac{1-\alpha}{\alpha}\left(\varepsilon C_2-\varepsilon C_1+(1-\varepsilon)C_1\right)>1, \tag{48}$$

that is, if

$$\alpha>\alpha^*\equiv\frac{C_1-\varepsilon\left(2C_1+C_2\right)}{1+C_2-C_1}.$$

Since $\alpha\downarrow 0$ violates (48) and $\alpha=1$ satisfies (48), there exists $\alpha^*\in(0,1)$, as desired.

The rest of the proof is the same as Part 3 of Proposition 8.