

Federal Reserve Bank of Minneapolis
Research Department Staff Report 430

July 2009

Reputation from Nested Activities: The Inefficient Effects of Scapegoating *

Guillermo L. Ordoñez

Federal Reserve Bank of Minneapolis and Yale University

ABSTRACT

Scapegoating is often said to be a source of inefficiency in organizations. In this paper, I analyze the consequences of scapegoating within a firm in a model where reputation concerns drive the actions of superiors. Consider delegation choices, for example. Hiring efficient workers may be a good idea if successful production is the only way to build reputation. But if successful scapegoating also increases reputation, superiors will tend to hire less efficient workers and will eventually blame them for failures.

I characterize scapegoating as an activity "nested" after failures. Even though the results of scapegoating do not affect welfare directly, they do so indirectly through the decisions governing the probability of success in production. We examine how activities "nested" after good results may increase efficiency without relying on costly incentives and why superiors tend to hire better workers during good times.

*E-mail: guillermo.ordonez@yale.edu. I would like to thank David K. Levine, Joe Ostroy, Larry Samuelson, Vasiliki Skreta, Jeroen Swinkels, Bill Zame, as well as seminar participants at UCLA, Minnesota, and Washington University in St. Louis, for useful comments. The usual waiver of liability applies. The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1 Introduction

A quick review of major newspapers from many countries shows that people commonly condemn scapegoating behavior. This attitude is due not only to its unfairness but also to its negative effects on efficiency and performance in organizations. In fact, this general view has been widely used to justify recent institutional reforms designed to improve efficiency by reducing superiors' scapegoating.¹ So far, however, no model has been developed to formalize this conventional and seemingly well-accepted wisdom.

In this paper, I study the impact of scapegoating on efficiency by focusing on delegation decisions made by reputation-concerned superiors in firms. We offer a novel interpretation of scapegoating as an irrelevant activity that only happens after failures and that may be used by reputation-concerned superiors as an additional way to signal their competence.

Given this characterization of scapegoating, I introduce the concept of "nested reputation" games as an environment where certain stages can be achieved only after certain situations occur. Even when "nested activities" may be irrelevant from the standpoint of consumers' welfare, they help to build superiors' reputation.² The existence of these activities will change not only superiors' incentives to make decisions but also the outcome of the game.

I introduce both delegation and scapegoating decisions into a reputation environment with adverse selection and moral hazard. In the model, superiors can be either competent or inept. These nonobservable types differ along two dimensions. First, competent have the choice to hire costly experts, while inept can only hire costless nonexperts, hence reducing their probability of producing successfully. Second, competent are also more effective than inept in blaming workers for failures. Hence, both production and scapegoating results are useful elements for consumers to infer the superior's type. When making their delegation decisions, competent superiors know that hiring experts increases the probability of good production results and gains in reputation. However they also know that hiring nonexperts can serve as a better buffer for reputation losses in case of failures.

I derive conditions for an efficient equilibrium with and without scapegoating. We conclude that it is more difficult, and sometimes even impossible, to achieve efficiency as an equilib-

¹The assignment of more responsibility to superiors has been a main goal of OECD institutional changes over the past decade. Examples include the Next Steps and Outcome-Output programs of the United Kingdom and New Zealand. Article 25 (RCSS) of the Rome Conference for an International Criminal Court also criticizes scapegoating from civilian superiors (Martin (1997); Polidano (1999)).

²To the best of my knowledge, nested reputation models do not exist. This work is an initial effort to understand how the results in a standard reputation game change when introducing "nested" activities as substitutes or complements to the original activity that generates reputation. Although this paper shares some features with the literature on "reputation spillover" (as in Cole and Kehoe (1996)), the logic is not the same. Spillovers deal with different and multiple types of reputation. Here, reputation is based on a single aspect, which is constructed through several nested stages and steps.

rium when allowing for scapegoating after failures. The main force driving this result is that hiring experts becomes less attractive when scapegoating is a possibility. First, scapegoating avoids a big decrease in reputation after a failure, reducing the expected gains from working with experts. Second, hiring experts hinders the use of scapegoating to maintain reputation after a failure (since it is harder to blame experts than nonexperts), reducing the expected reputation losses from working with nonexperts.

Even when focusing on how scapegoating, an activity after failures, negatively affects efficiency, we also extend the logic to study how irrelevant activities that only occur after successes increase the probabilities of achieving efficiency by exploiting reputation forces in the right direction, without requiring monetary resources or costly incentives. Many examples, from areas as diverse as sports and education, are discussed in the paper.

Finally, I show that considering reputation from nested activities allows us to identify a "Machiavellian effect,"³ which can be stated as "Superiors tend to hire nonexperts in bad times and experts in good times."

The literature on reputation is very large,⁴ but the literature on scapegoating is almost nonexistent. Despite the recognition of strategic reasons for scapegoating in the social-psychology literature (Bell and Tetlock (1989); Douglas (1995)), formal economic studies of this behavior are new and sparse. Dezsó (2004) analyzes the conditions under which random firing of potential innocents (scapegoats) is a reaction to failures in order to maintain reputation. He focuses on firing and not on hiring, without being able to analyze the impact on efficiency. Segendorff (2000) analyzes the possible hiring of scapegoats using a signaling game, also without analyzing efficiency consequences. Winter (2001) finds that, under some circumstances, in order to provide better incentives to top levels in an organization, it may be optimal for middle levels to bear more responsibility, an aspect he labels "scapegoating." He does not consider reputation effects nor hiring decisions, however.⁵

In Section 2, I discuss the basic model of reputation with delegation and scapegoating, a special application of the proposed "nested reputation" environment. In Section 3, I analyze the conditions for the existence of an efficient equilibrium and show how and when scapegoating

³In his famous book *The Prince*, Machiavelli wrote, "Princes should delegate to others the enactment of unpopular measures and keep in their own hands the distribution of favours." Machiavelli's argument was that princes should delegate when the probability of having a good outcome is low and work by themselves if it is high. In this way, princes would be able to blame others if something goes wrong, thereby maintaining their reputation. More recently, Alesina and Tabellini (2005, 2007, 2008) formally modeled this pattern among politicians.

⁴Starting with Kreps and Wilson (1982) and Milgrom and Roberts (1982), important contributions on reputation models are Fudenberg and Levine (1989, 1992), Mailath and Samuelson (2001), and Cripps, Mailath, and Samuelson (2004).

⁵Empirical studies about scapegoating are even less common. An exception is Huson, Malatesta, and Parrino (2004), who developed a moral hazard driven scapegoat hypothesis based on agency models to study the impact of managerial succession on firm performance.

leads to inefficiency and to the “Machiavellian effect.” In Section 4, I propose a way to induce efficiency by exploiting reputation concerns, without relying on the use of costly incentives. In Section 5, I make some final remarks.

2 The Model

2.1 Description

This model extends Mailath and Samuelson (2001) by introducing delegation and scapegoating as a nested activity. Assume a superior (the president of a country, a minister, the owner of a firm, or a CEO) who is responsible for providing a service, selling a good, or, in general achieving a target that generates utility to “consumers” (who can also be citizens, stockholders, or even upper-level superiors in the hierarchy).

Each period the superior has to make an unobservable delegation decision to achieve the target. He can be one of two possible types, competent (C) or inept (I). Competents have two possible choices: to hire experts (E), paying a wage $w > 0$, or to hire nonexperts (N), paying 0. Inepts can only delegate to nonexperts (N).⁶ We can rationalize the existence of these two types in several ways. The simplest one is to assume that workers are able to observe the superior’s type and that experts do not work with inepts (who may generate some disutility, such as difficulties for professional development). Another possibility is to assume that competent have the skills to perfectly identify who is an expert and who is a nonexpert, while inepts do not have access to this screening (or interview) technology.⁷

Before deciding, superiors observe the state of the nature, good (G) or bad (B), which affects the probability of success in achieving the target. After deciding, a nondeterministic output is obtained, which can be good (g) or bad (b). When competent hire experts, the probability of a good result in good times is $(1 - \rho) > \frac{1}{2}$ and in bad times $\alpha < \frac{1}{2}$. When superiors hire nonexperts, the probability of a good result in good times is $(1 - \alpha)$ and in bad times ρ .⁸ I assume $(\alpha - \rho) > w > 0$; hence, it is efficient to society that competent always hire experts, both in good times and in bad. If “consumers” knew agents’ delegation decisions, they would be willing to pay a premium for competent to always hire experts.

⁶Employees in this model are just dummies who do not make any decisions. Hence, to assume the rent of machines instead of delegation does not change the analysis.

⁷In the latter case, inepts may prefer to hire employees at random (by offering a wage w) if the proportion of experts in the workers’ population is high enough. Even in the case where they eventually hire experts by following this strategy, the model’s conclusions remain unchanged, provided that some positive proportion of nonexperts exists (the reasons will be clarified later while discussing the model). Hence, for expositional purposes, we will assume that inepts do not have the possibility to attract experts.

⁸The assumption of symmetry in probabilities does not change the main conclusions but allows the use of just two parameters (α and ρ) rather than four, thus eliminating awkward expressions.

If the outcome is a failure, the superior may have the opportunity to create a report about its causes, which implies that a decision must be made regarding the intensity and amount of evidence to be presented against workers (scapegoating). This is a nested second stage in the game that occurs only after failures in production, not after successes. Once the report is done, a nondeterministic decision about the credibility of the evidence is taken, by a "court", for example, that concludes whether the employee (e_c) or the superior (s_c) has to be considered the culprit of the failure.⁹

By deciding the blaming intensity and the amount of evidence displayed, superiors directly choose the probability that the court will blame the worker.¹⁰ In particular, inepts choose a probability x the court will pronounce against subordinates such that $x \in [0, \bar{x}]$ where $\bar{x} \leq 1$. Hence, maximum blaming intensities do not necessarily guarantee that the court will decide against subordinates. When competents choose a probability that the court will decide in his favor, they know whether the blamed employee is an expert or a nonexpert. If competents worked with experts, they choose probability y and if they worked with nonexperts they may choose a different one, z . These probabilities are $y \in [0, \bar{y}]$ and $z \in [0, \bar{z}]$, where $\bar{y} \leq \bar{z} \leq 1$. The blaming capabilities under maximum blaming intensities (\bar{x} , \bar{y} and \bar{z}) are known exogenous parameters.

Finally, at the end of the period, the superior may be replaced by another superior with a fixed probability λ . The substitute will be competent with a probability $\theta \in (0, 1)$.¹¹

"Consumers" repeatedly receive the output generated under superiors' commands (e.g, consumers purchase a good, citizens receive a service, and stockholders obtain dividends).¹² This generates two possible utility levels in each period: 1 if the result is a good outcome ($u(g) = 1$) and 0 if it is a bad outcome ($u(b) = 0$). Consumers do not get any utility from scapegoating.

Consumers do not observe the superiors' type and delegation decisions and cannot see the state of the economy. They know only the probability of being in a good state is $\Pr(G) = \gamma$ and they observe only the results from production activities (success or failure) and from scapegoating activities after failures (superior or employee considered culprit).¹³ Using this information, consumers update the probability that the superior is competent, $\Pr(C) = \phi$.

⁹When referring to a "court," I am thinking not only of a judiciary court but also a court of public opinion, a board of directors, or in general any group that decides about the assignment of responsibility by considering the existing evidence.

¹⁰No blaming intensity and no evidence, for example, makes it impossible for the court to decide against the worker. Increasing blaming efforts also increases the probability that the court will pronounce against employees.

¹¹This assumption is needed to sustain an efficient equilibrium in the long run, as discussed in Mailath and Samuelson (2001) and Cripps, Mailath, and Samuelson (2007).

¹²By "consumers", we mean in general a continuum of identical individuals of unit mass (such that no single individual can affect the future play of the game) who obtain a good or service from the superior.

¹³All consumers receive the same public result or signal.

This is important to superiors, since we assume that each consumer has to buy the good or service before production takes place, hence, paying the expected utility and not the real utility it delivers. The greater the reputation (probability of the superior being competent ϕ), the greater the probability assigned by consumers to obtain good outcomes and the more payments they will be willing to make for the good or service. This is why superiors are so concerned about reputation, while consumers are concerned only about the utility derived from production.

2.2 Timing

The timing of the model is as follows:

0) The superior receives the payment for period t , before the production takes place, which only depends on his reputation ϕ at the beginning of t , and not on his period t 's true type, delegation decision, or production result.

1) The superior observes ϕ , w , and the environment state (G or B). Competents decide to hire experts or nonexperts. Inepts can only attract and hire nonexperts. Consumers do not observe this decision, nor do they observe whether there are good or bad times.

2) Output is produced, and both consumers and the superior observe the true utility given by a good (g) or bad (b) outcome (1 or 0, respectively). All consumers receive the same public realization of utility outcome.

3) The superior has to report the cause of the failure in case of a bad outcome, deciding blaming intensities and how much evidence to present against employees (i.e., x , y , or z depending on the type of superior and employee).

4) A court decides if the employee was the culprit (e_c) or if the superior was the culprit (s_c) of the failure.

5) With probability λ the superior is replaced by another one, who is competent with a probability θ .

2.3 Equilibrium Definition

First, we restrict behavior to be Markov in order to eliminate equilibria that depend on implausible degrees of coordination between the superior's behavior and consumers beliefs about that superior's behavior.¹⁴ Hence, under uncertainty about the superior's type, the

¹⁴Also, by restricting attention to Markov strategies, in equilibrium different superiors behave identically in identical situations. See discussions in Mailath and Samuelson (1998, 2006).

state variable is the probability assigned by consumers to the superior being competent (i.e., the reputation denoted as ϕ).

Before production, a Markov strategy for competents is a mapping $\tau_k : [0, 1] \rightarrow [0, 1]$, where $\tau_k(\phi)$ is the probability of hiring an expert when reputation is $\phi \in [0, 1]$ and the state of nature is $k \in \{B, G\}$. Inepts make no choice, having then a trivial strategy of hiring nonexperts.

After a bad result in production, a Markov strategy for competents that hired experts is a mapping $y : [0, 1] \rightarrow [0, \bar{y}]$, where $y(\phi)$ is the probability that the court will decide against the employee when reputation is $\phi \in [0, 1]$. I will refer to this strategy as blaming intensities. The same strategy is available for inepts ($x(\phi)$) and competents who hired nonexperts ($z(\phi)$).

The behavior of consumers is described by the Markov belief function $p : [0, 1] \rightarrow [0, 1]$, where $p(\phi)$ is the probability that consumers assign to receiving a good outcome, given a reputation $\phi \in [0, 1]$ (recall that utilities from good and bad results have been normalized to 1).

In a Markov perfect equilibrium, superiors maximize expected discounted profits, consumers beliefs about superiors' actions are correct, and they use Bayes' rule when possible to update their inference about superiors' types. There are possible two rounds of updating: the update after production results ($\Pr(C|g)$ and $\Pr(C|b)$) and the potential update after scapegoating results ($\Pr(C|b, e_c)$ and $\Pr(C|b, s_c)$). As an example, $\Pr(C|b)$ can be written explicitly in terms of parameters and decision rules as

$$\Pr(C|b) = \frac{\Pr(b|C)\phi}{\Pr(b|C)\phi + \Pr(b|I)(1 - \phi)}, \quad (1)$$

where

$$\begin{aligned} \Pr(b|C) &= \gamma[\rho\tau_G + \alpha(1 - \tau_G)] + (1 - \gamma)[(1 - \alpha)\tau_B + (1 - \rho)(1 - \tau_B)] \\ \Pr(b|I) &= \gamma\alpha + (1 - \gamma)(1 - \rho). \end{aligned}$$

Before defining the equilibrium, we need to define the value function for the superior as a function of the reputation value ϕ . For competents,

$$V_k(\phi) = \max_{\tau_k, y, z} \{p(\phi) - \tau_k w + \delta(1 - \lambda)E[V(\phi')|\tau_k, y, z]\}, \quad (2)$$

where $k = \{B, G\}$ and expectation is constructed over possible states of nature and possible reputation levels next period (ϕ' , which is a function of delegation and blaming decisions).

For inepts,

$$V(\phi) = \max_x \{p(\phi) + \delta(1 - \lambda)E[V(\phi')|x]\}. \quad (3)$$

Definition 1 A Markov perfect equilibrium is: probabilities of hiring experts in both good and bad times ($\tau_G(\phi)$ and $\tau_B(\phi)$), blaming intensities ($y(\phi)$, $x(\phi)$ and $z(\phi)$), probabilities consumers assign to receiving a good outcome ($p(\phi)$) given a reputation prior $\phi = \Pr(C)$, and posterior beliefs $\varphi = \Pr(C|R, \phi)$, where R are the three possible results $R \in \{g; (b, e_c); (b, s_c)\}$, such that

1) Delegation decisions by competent superiors

$\tau_G(\phi)$ (in good times) and $\tau_B(\phi)$ (in bad times) maximize the value function $V(\phi)$ (eq. 2) for all possible reputation values ϕ

2) Blaming intensities by superiors

$x(\phi)$, $y(\phi)$ and $z(\phi)$ maximize the value function $V(\phi)$ (eq. 2 and 3) for all feasible ϕ

3) Expected utility (and payments) of consumers

Probabilities consumers assign to receiving a good outcome given a reputation prior ϕ (i.e., profits for the superior)

$$p(\phi) = \Pr(g|\phi) = \Pr(g|C)\phi + \Pr(g|I)(1 - \phi) \quad (4)$$

4) Beliefs about competence (updated using Bayes' rule)

a) Update after a good outcome (g)

$$\varphi(\phi|g) = \phi_g = (1 - \lambda) \Pr(C|g) + \lambda\theta \quad (5)$$

b) Update after the court considers the employee responsible for a bad outcome (b, e_c)

$$\varphi(\phi|b, e_c) = \phi_b^{e_c} = (1 - \lambda) \Pr(C|b, e_c) + \lambda\theta \quad (6)$$

c) Update after the court considers the superior responsible for a bad outcome (b, s_c)

$$\varphi(\phi|b, s_c) = \phi_b^{s_c} = (1 - \lambda) \Pr(C|b, s_c) + \lambda\theta \quad (7)$$

5) Principals' beliefs about the agent's actions are correct.

A strategy for agents uniquely determines the equilibrium updating rule that principals must use if their beliefs are to be correct.

3 Efficient Equilibrium, Inefficient Scapegoating

A fundamental question in this paper is: Does scapegoating really reduce the probability of achieving an efficient outcome? Hence, we focus on the conditions for an efficient situation

to be sustained as an equilibrium.¹⁵ Considering the assumption $(\alpha - \rho) > w > 0$, efficiency is achieved when competents always hire experts, regardless of their current reputation or whether times are good or bad (i.e., $\tau_G(\phi) = \tau_B(\phi) = 1$, for all feasible ϕ).

The condition for this efficient situation to be sustained as an equilibrium is expressed by a cutoff Δ , such that wages w have to be smaller than Δ . This cutoff is obtained in both good and bad times with scapegoating possibilities (Δ_G^S and Δ_B^S) and without scapegoating possibilities (Δ_G^{NS} and Δ_B^{NS}). The last case is used as a benchmark to see how results differ when superiors are allowed to explain bad results, potentially blaming workers. Whenever $\bar{y} > \bar{x}$ and scapegoating abilities are high enough (specifically when a sufficient condition $\bar{z} \geq 1 - \frac{\rho}{\alpha}(1 - \bar{y})$ holds),

$$\Delta_G^{NS} = \Delta_B^{NS} \geq \Delta_G^S \geq \Delta_B^S. \quad (8)$$

Given wages in the economy, these simple inequalities, which are in fact typically strict, summarize the main conclusions of the paper. The first inequality means that scapegoating makes the condition for an efficient equilibrium $\Delta \geq w > 0$ more difficult to hold. Furthermore, since $\Delta_G^{NS} = \Delta_B^{NS} > 0$, in the absence of scapegoating it is always possible to find a positive wage differential that sustains efficiency, which is not necessarily the case with scapegoating. The second inequality means that it is even more difficult to achieve efficiency with scapegoating in bad times rather than in good times (the Machiavellian effect).

3.1 Conditions for Efficient Equilibrium

As a first step, we present the condition for the existence of an efficient equilibrium without scapegoating. In this case, only two possible states (g and b) are possible, since there is no blaming activity allowed after a failure (*nobody asks why things went wrong!*). The reputation after a bad draw would be ϕ_b directly. In a similar vein, the reputation after two consecutive bad results ($\varphi(\varphi(\phi|b)|b)$) will be denoted as ϕ_{bb} . The proof is in the Appendix.

Proposition 2 *Efficient Equilibrium without Scapegoating*

Suppose $\lambda \in (0, 1)$, $\phi_0 \in [\lambda\theta, 1 - \lambda(1 - \theta)]$, $\delta \in (0, 1)$, and $\theta \in (0, 1)$. If the report about the causes of the failure is not allowed (no "blaming" stage), then in both good and bad times there exists a positive cutoff

$$\Delta^{NS} = \Delta_G^{NS} = \Delta_B^{NS} = \min_{\phi \in [\lambda\theta, 1 - \lambda(1 - \theta)]} \{\delta(1 - \lambda)[X + \delta(1 - \lambda)V_f]\} > 0 \quad (9)$$

¹⁵This model has multiple equilibria, including a very inefficient one that may arise without conditions, in which competents hire only nonexperts. Intuitively, in this case, consumers do not have statistical elements to update their beliefs about the superior's type. Since there are no gains in terms of reputation from hiring experts, competents optimally prefer to hire nonexperts rather than experts and pay higher wages.

such that, for all wages $\Delta^{NS} \geq w > 0$, the efficient pure strategy profile in which competents always hire experts is a Markov perfect equilibrium,

where

$$\begin{aligned} V_f &= \Pr(g|E)Y_g + \Pr(b|E)Y_b \\ X &= (\alpha - \rho)[p(\phi_g) - p(\phi_b)] \\ Y_i &= (\alpha - \rho)[V(\phi_{gi}) - V(\phi_{bi})] \quad \text{for } i \in \{g, b\} \end{aligned}$$

with $V(\phi)$ defined in equation (2).

Since our objective is to compare this benchmark with the extended model with scapegoating, the next proposition shows conditions to have an efficient equilibrium when blaming is a possibility (people ask why things went wrong!). The proof is in the Appendix.

Proposition 3 Efficient Equilibrium with Scapegoating

Suppose $\lambda \in (0, 1)$, $\phi_0 \in [\lambda\theta, 1 - \lambda(1 - \theta)]$, $\delta \in (0, 1)$, and $\theta \in (0, 1)$. If the report about the causes of the failure is allowed (scapegoating),

a) If $\bar{y} \leq \bar{x}$, conditions for an efficient equilibrium are exactly the same as the case without scapegoating (Proposition 2).

b) If $\bar{y} > \bar{x}$, there exists a, not necessarily positive, cutoff for each state of the world $k \in \{B, G\}$

$$\Delta_k^S = \min_{\phi \in [\lambda\theta, 1 - \lambda(1 - \theta)]} \left\{ \delta(1 - \lambda)[X^k + \delta(1 - \lambda)V_f^k] \right\} \quad (10)$$

such that, for all wages $\Delta_k^S \geq w > 0$, the efficient pure strategy profile in which competents always hire experts is a Markov perfect equilibrium,

where

$$\begin{aligned} V_f^k &= \Pr(g|E)Y_g^k + \Pr(b|E)Y_b^k \\ X^B &= (\alpha - \rho)p(\phi_g) + (1 - \alpha)p(\phi_{b,E}) - (1 - \rho)p(\phi_{b,N}) \\ X^G &= (\alpha - \rho)p(\phi_g) + \rho p(\phi_{b,E}) - \alpha p(\phi_{b,N}) \end{aligned}$$

being

$$\phi_{b,E} = \bar{y}\phi_b^{ec} + (1 - \bar{y})\phi_b^{sc} \quad (11)$$

$$\phi_{b,N} = \bar{z}\phi_b^{ec} + (1 - \bar{z})\phi_b^{sc} \quad (12)$$

and, for $i \in \{g, b\}$,

$$Y_i^B = (\alpha - \rho)V(\phi_{gi}) + (1 - \alpha)[\bar{y}V(\phi_{bi}^{ec}) + (1 - \bar{y})V(\phi_{bi}^{sc})] - (1 - \rho)[\bar{z}V(\phi_{bi}^{ec}) + (1 - \bar{z})V(\phi_{bi}^{sc})]$$

$$Y_i^G = (\alpha - \rho)V(\phi_{gi}) + \rho[\bar{y}V(\phi_{bi}^{ec}) + (1 - \bar{y})V(\phi_{bi}^{sc})] - \alpha[\bar{z}V(\phi_{bi}^{ec}) + (1 - \bar{z})V(\phi_{bi}^{sc})].$$

Before going to the main proposition of the paper, it is important to emphasize two features of the solution. First, the non scapegoating case is just a particular example of the scapegoating case. If $\bar{y} \leq \bar{x}$, both cases are in fact exactly the same. If $\bar{y} > \bar{x}$, as $\bar{x}, \bar{y}, \bar{z} \rightarrow 0$ (maintaining the relation $\bar{z} \geq \bar{y} > \bar{x}$) always $\phi_{b,N} \rightarrow \phi_{b,E} \rightarrow \phi_b$ (as can be checked easily from equations (6), (7), (11) and (12)). Hence, $X^B \rightarrow X^G \rightarrow X = (\alpha - \rho)[p(\phi_g) - p(\phi_b)]$, and $V_f^B \rightarrow V_f^G \rightarrow V_f$. This is equivalent to stating that all cutoffs approach ($\Delta_B^S \rightarrow \Delta_G^S \rightarrow \Delta^{NS}$), or that Proposition 3 approaches Proposition 2 as the importance of blaming disappears.

Second, in the differing case ($\bar{y} > \bar{x}$), since w is positive by assumption and there is no way to know the sign of Δ_k^S , it can only be said that whenever $\Delta_k^S < 0$, no wage can possibly support an efficient equilibrium. Even though without scapegoating there is always a positive wage that supports an efficient equilibrium, this is not necessarily the case with scapegoating. This implies that scapegoating is harmful for efficiency, which will be formalized in the following subsections.

In the remainder of the paper, and unless stated otherwise, when referring to the scapegoating case, I refer specifically to the case in which blaming makes a difference ($\bar{y} > \bar{x}$).

3.2 Scapegoating Inefficiency

Here, I show the difficulties scapegoating imposes in achieving efficiency. I do this by proving that the conditions for efficiency with scapegoating (when $\bar{y} > \bar{x}$ from Proposition 3) are more difficult to hold than the conditions for efficiency without scapegoating (from Proposition 2).

Proposition 4 *Scapegoating Inefficiency*

Suppose $\lambda \in (0, 1)$, $\phi_0 \in [\lambda\theta, 1 - \lambda(1 - \theta)]$, $\delta \in (0, 1)$, $\theta \in (0, 1)$ and competents have more blaming abilities than inepts ($\bar{y} > \bar{x}$). It is always possible to find a $\bar{z} \geq z^ = 1 - \frac{\rho}{\alpha}(1 - \bar{y})$ such that the range of wages $w > 0$ that supports an efficient situation is smaller with scapegoating than without it.*

Proof. We need to prove that $\Delta^{NS} \geq \Delta_G^S$ for all $\phi \in (0, 1)$. This is a sufficient condition since the Machiavellian effect theorem discussed in Proposition 7 below shows that, in all

cases, $\Delta_G^S \geq \Delta_B^S$. This proof is based on the simpler case in which scapegoating is possible only in the current period, not in the future. The conclusion for the more general case is the same, but it is characterized by awkward statements (shown in the Appendix). I consider only the relevant case in which $\bar{y} > \bar{x}$ and there is a separating blaming equilibrium such that $\phi_b^{ec} > \phi_b > \phi_b^{sc}$.

I proceed in three steps. First, I show that $\phi_{b,N} \geq \phi_{b,E}$, second, that $\phi_{b,E} \geq \phi_b$ (as defined in Proposition 3), and finally, that $\Delta^{NS} \geq \Delta_G^S$ by proving that $X + \delta(1-\lambda)V_f \geq X^G + \delta(1-\lambda)V_f^G$ for all feasible ϕ .

Step 1: ($\phi_{b,N} \geq \phi_{b,E}$)

Consider beliefs about decision rules in the efficient equilibrium ($\tau_G(\phi) = \tau_B(\phi) = 1$), from equations (12), (6), and (7),

$$\phi_{b,N} = \bar{z}\phi_b^{ec} + (1-\bar{z})\phi_b^{sc}$$

$$\phi_{b,N} = (1-\lambda) \left[\bar{z} \frac{\bar{y}\phi_b^{pb}}{\bar{y}\phi_b^{pb} + \bar{x}(1-\phi_b^{pb})} + (1-\bar{z}) \frac{(1-\bar{y})\phi_b^{pb}}{(1-\bar{y})\phi_b^{pb} + (1-\bar{x})(1-\phi_b^{pb})} \right] + \lambda\theta$$

and, from equations (11), (6), and (7),

$$\phi_{b,E} = \bar{y}\phi_b^{ec} + (1-\bar{y})\phi_b^{sc}$$

$$\phi_{b,E} = (1-\lambda) \left[\bar{y} \frac{\bar{y}\phi_b^{pb}}{\bar{y}\phi_b^{pb} + \bar{x}(1-\phi_b^{pb})} + (1-\bar{y}) \frac{(1-\bar{y})\phi_b^{pb}}{(1-\bar{y})\phi_b^{pb} + (1-\bar{x})(1-\phi_b^{pb})} \right] + \lambda\theta,$$

where $\phi_b^{pb} = \Pr(C|b)$.¹⁶

Subtract both expressions:

$$\phi_{b,N} - \phi_{b,E} = \frac{(1-\lambda)\phi_b^{pb}(1-\phi_b^{pb})(\bar{z}-\bar{y})(\bar{y}-\bar{x})}{\bar{x}(1-\bar{x}) + \phi_b^{pb}(\bar{y}-\bar{x})[1-2\bar{x} - \phi_b^{pb}(\bar{y}-\bar{x})]}, \quad (13)$$

which cannot be negative, since $\bar{z} \geq \bar{y} > \bar{x}$ and $\phi_b^{pb} \in [0, 1]$.

Step 2: ($\phi_{b,E} \geq \phi_b$)

Subtract $\phi_b = (1-\lambda)\Pr(C|b) + \lambda\theta$ from equation (11):

$$\phi_{b,E} - \phi_b = (1-\lambda) \left[\bar{y} \frac{\bar{y}\phi_b^{pb}}{\bar{y}\phi_b^{pb} + \bar{x}(1-\phi_b^{pb})} + (1-\bar{y}) \frac{(1-\bar{y})\phi_b^{pb}}{(1-\bar{y})\phi_b^{pb} + (1-\bar{x})(1-\phi_b^{pb})} - \phi_b^{pb} \right],$$

¹⁶Recall $\phi_b^{pb} = \Pr(C|b)$ represents the standard Bayes updating after a bad outcome and before any blaming activity (the superscript pb denotes "pre-blaming"). This is an update not adjusted by λ because it happens before the period ends and a replacement occurs.

which implies that

$$\phi_{b,E} - \phi_b = \frac{(1 - \lambda)\phi_b^{pb}(1 - \phi_b^{pb})^2(\bar{y} - \bar{x})^2}{\bar{x}(1 - \bar{x}) + \phi_b^{pb}(\bar{y} - \bar{x})[1 - 2\bar{x} - \phi_b^{pb}(\bar{y} - \bar{x})]}, \quad (14)$$

which cannot be negative, since $\phi_b^{pb} \in [0, 1]$.

Step 3: ($\Delta^{NS} \geq \Delta_G^S$ for all $\phi \in (0, 1)$)

By equations (9) and (10), it is sufficient to show the following two claims.

Claim 1) $X \geq X^G$ for all ϕ .

Subtract these expressions:

$$X - X^G = \alpha[p(\phi_{b,N}) - p(\phi_b)] - \rho[p(\phi_{b,E}) - p(\phi_b)],$$

which is non-negative, since $\alpha > \rho$ by assumption and $p(\phi_{b,N}) \geq p(\phi_{b,E})$ for all feasible ϕ by step 1 (equation 13) and monotonicity of $p(\phi)$.

Claim 2) $V_f \geq V_f^G$ for all ϕ .

Subtract these expressions:

$$V_f - V_f^G = \Pr(g|E)[Y_g - Y_g^G] + \Pr(b|E)[Y_b - Y_b^G],$$

which is non-negative if, for $i \in \{g, b\}$,

$$Y_i - Y_i^G = (\alpha\bar{z} - \rho\bar{y})[V(\phi_{bi}^{ec}) - V(\phi_{bi}^{sc})] - (\alpha - \rho)[V(\phi_{bi}) - V(\phi_{bi}^{sc})] \geq 0.$$

Because of the monotonicity of $V(\phi)$ in ϕ and since, by equation (6), $\phi_b^{ec} \geq \phi_b$, then $V(\phi_{bi}^{ec}) \geq V(\phi_{bi})$. A sufficient condition for non-negativity is then $(\alpha\bar{z} - \rho\bar{y}) \geq (\alpha - \rho)$, which is the same as

$$\bar{z} \geq z^* = 1 - \frac{\rho}{\alpha}(1 - \bar{y}), \quad (15)$$

where $\frac{\rho}{\alpha}$ is a measure of the relative capability of experts to achieve good production results when compared to nonexperts.

Hence, whenever the sufficient condition $\bar{z} \geq z^*$ holds, regardless of the value function, the likelihood of having an efficient situation reaches its maximum without scapegoating. ■

The intuition behind this result is that reports after a failure when $\bar{y} > \bar{x}$ represent a way for competents to further signal their competence. If this is the case, competents can exploit dif-

ferences in blaming abilities as an additional channel to distinguish themselves from inepts. This can be done in two ways: first, by the difference between competents and inepts in maximum blaming capabilities ($\bar{y} - \bar{x}$), which reduces reputation losses after bad results: and second, by the assumed difference between blaming experts and nonexperts ($\bar{z} - \bar{y}$), which may introduce an additional gain from hiring nonexperts (additional to low wages) by increasing the probability that the court assigns the responsibility of the failure to the employee.

To interpret the sufficient condition for the inefficiency of scapegoating, $z \geq z^*$, recall that z^* depends on $\frac{\rho}{\alpha}$, a measure of the relative abilities of experts to achieve good results in production when compared to nonexperts. If hiring experts almost guarantees success in good times ($\rho \rightarrow 0$), then $z^* \rightarrow 1$. If hiring experts does not significantly increase the probability of success ($\rho \rightarrow \alpha$), then $z^* \rightarrow \bar{y}$. Hence, the sufficient condition $\bar{z} \geq z^*$ is more difficult to hold when hiring experts is really beneficial from a productive point of view and superiors can signal their competence directly in the first stage. Contrarily, if hiring experts does not make an important difference in production, competents tend to rely more on the use of scapegoating to signal competence, thus contributing to inefficiency.¹⁷

Conditions for an efficient equilibrium in the three cases discussed previously, without scapegoating (both in good times and in bad, Δ^{NS}), with scapegoating in good times (Δ_G^S), and with scapegoating in bad times (Δ_B^S) are exemplified in Figure 1, where I assumed that $V_f = V_f^G = V_f^B = 0$. This assumption conservatively biases the results in favor of hiring experts. In this conservative example, banning scapegoating (*by not asking why things went wrong!*) increases the incentives to efficiently hire experts.¹⁸

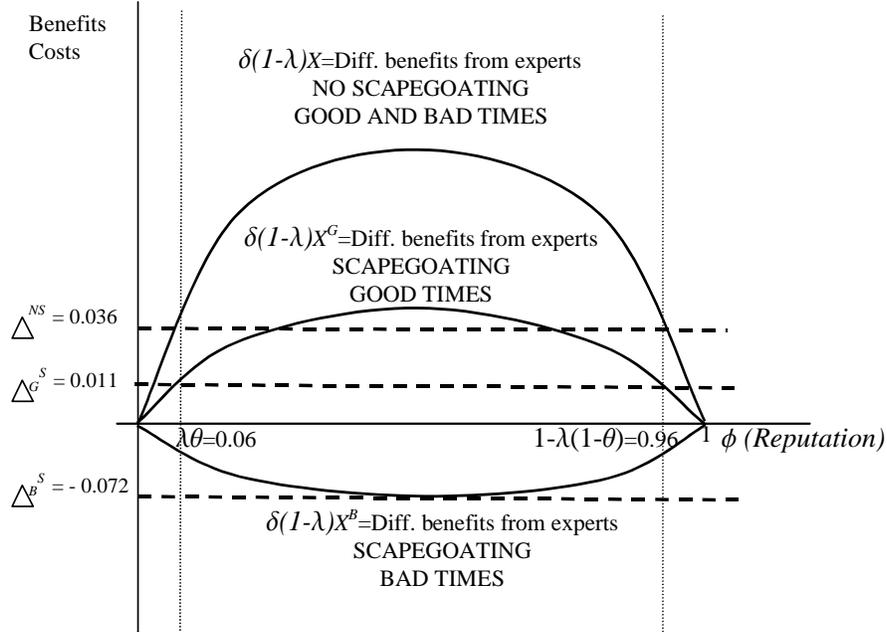
3.3 The Intuition Behind the Inefficiency of Scapegoating

The importance of blaming for efficiency resides both in its value to competents to further signal their competence and in its irrelevance to consumers. By assumption, superiors care only about reputation, and their decisions react more to activities that better maintain or construct reputation. If those activities occur only after particular situations, such as scapegoating that happens only after failures, superiors make decisions based on trying to facilitate those situations to arise, even when they are detrimental to activities that really matter to consumers. This effect will be even more important in cases in which superiors value reputation beyond the impact on profits (for example, reputation as a career booster).

¹⁷The sufficient condition is relevant only because we do not know the shape of the value function. However, if the value function is linear (risk neutrality) or convex on ϕ , scapegoating would always imply inefficiency, regardless of the specific value of \bar{z} .

¹⁸Parameters used: $\lambda = 0.1$, $\theta = 0.6$, $\delta = 0.99$, $\rho = 0.1$, $\alpha = 0.4$, $\gamma = 0.5$, $\bar{x} = 0.15$, $\bar{y} = 0.3$, and $\bar{z} = 0.85$. The sufficient condition from equation (15) holds because, in this case, $\bar{z} > z^* = 0.825$.

Figure 1: Example of Conditions for Efficient Equilibrium



Assume, for example, the extreme case in which $\bar{x} = 0$ and $\bar{y} = \bar{z} = 1$ (i.e., inepts cannot convince anybody about the blame of employees, while competents can always blame convincingly). In this situation, if consumers see that after a failure the court decides against subordinates, they clearly learn that the superior is competent. Here, blaming is better than production for competents to construct reputation. In fact, competents prefer to have a failure to signal their type more effectively through the court rather than through production performance. In this extreme example, competents never hire experts, since nonexperts are not only costless but also increase the probability of going to court.

Naturally, the previous extreme example is consistent with the case $\bar{y} > \bar{x}$. But what happens if $\bar{y} \leq \bar{x}$? As shown formally, in this case blaming does not have any effect on efficiency conditions. This result is a version of a cheap talk game. Inepts can always pool with competents' blaming strategies, adjusting the blaming intensity downward (not presenting proofs, burning evidence against employees, etc.).¹⁹ Since competents can always be imitated by inepts, there is no equilibrium where the blaming report can be used to further change reputation.

The whole action in previous propositions and proofs comes from the comparison of reputation that competents expect to obtain from hiring experts as opposed to nonexperts. Without scapegoating, the reputation conditional on the first round's results is known and given by ϕ_g

¹⁹Blaming is an activity, as are many others, where the success probability can be easily adjusted downward (due to laziness) but not upward.

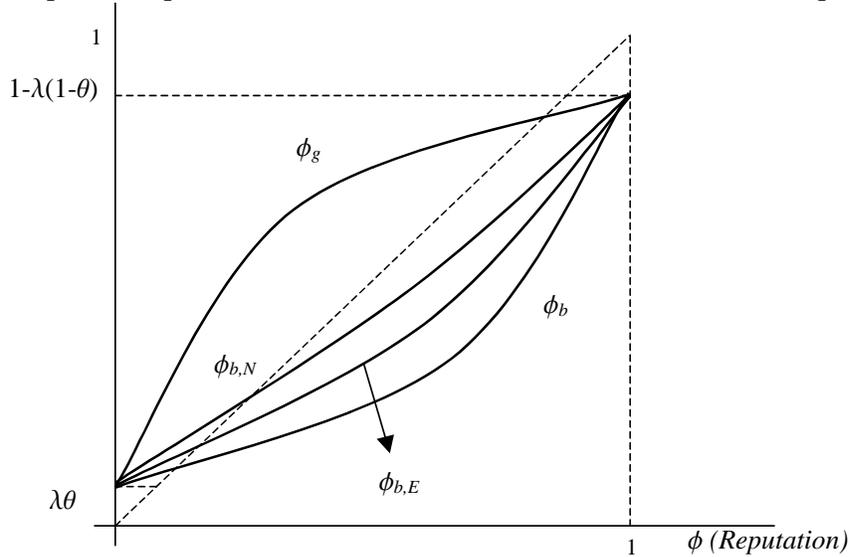
after a success and ϕ_b after a failure. With scapegoating, while the expected reputation after a good outcome is also independent of the hiring decision, ϕ_g (because the game ends after a success), the expected reputation after bad results depends on the hiring decision (since blaming nonexperts is easier). In particular, the expected reputation after a failure from hiring experts is $\phi_{b,E}$ (equation (11)) and from hiring nonexperts is $\phi_{b,N}$ (equation (12)).

The differential gains in reputation expected from having good production outcomes determine the incentives to hire experts. These gains are given by $(\phi_g - \phi_b)$ without scapegoating (regardless of the hiring decision), by $(\phi_g - \phi_{b,E})$ with scapegoating if hiring experts, and by $(\phi_g - \phi_{b,N})$ with scapegoating if hiring nonexperts. Hence, to understand incentives for expert hiring we need to understand how $\phi_{b,N}$, $\phi_{b,E}$, and ϕ_b relate to each other. As shown in steps 1 and 2 of the proof for Proposition 7, when $\bar{z} > \bar{y} > \bar{x}$, there is a clear ordering between these expressions.

$$\phi_g > \phi_{b,N} > \phi_{b,E} > \phi_b \quad (16)$$

Graphically, Figure 2 shows the reputation posterior for each case and possible prior ϕ . As can be seen, the updates for all possible beliefs are equal when ϕ is either zero or one. If the prior is $\phi = 0$, the update in all cases is $\lambda\theta$. If the prior is $\phi = 1$, the update is $1 - \lambda(1 - \theta)$. For all other values $\phi \in (0, 1)$, reputation updates have the ordering given by equation (16).

Figure 2: Expected reputation after the first round (with and without scapegoating)



The difference $(\phi_{b,E} - \phi_b)$ in equation (13) can be interpreted as the reduction in the incentives to hire experts, and the difference $(\phi_{b,N} - \phi_{b,E})$ in equation (14) can be seen as the increase in the incentives to hire nonexperts. While the expression $(\bar{y} - \bar{x})$ in the numerator of equations (13) and (14) shows the magnitude of reputation maintenance due to scapegoating, the ex-

pression $(\bar{z} - \bar{y})$ in the numerator of (13) shows the additional benefits from hiring nonexperts by taking advantage of scapegoating after a failure.

The following lemmas show that both $(\phi_{b,N} - \phi_{b,E})$ and $(\phi_{b,E} - \phi_b)$ not only are positive (as shown in steps 1 and 2 of the proof for Proposition 4) but also depend positively on gaps in the blaming abilities.

Lemma 5 *The difference between expected reputation after a failure from hiring experts versus hiring nonexperts $(\phi_{b,N} - \phi_{b,E})$ is non decreasing in both $(\bar{z} - \bar{y})$ and $(\bar{y} - \bar{x})$.*

Proof. Taking the derivative of expression $(\phi_{b,N} - \phi_{b,E})$ in equation (13) with respect to $(\bar{z} - \bar{y})$ and with respect to $(\bar{y} - \bar{x})$,

$$\frac{\partial[\phi_{b,N} - \phi_{b,E}]}{\partial(\bar{z} - \bar{y})} = \frac{(1 - \lambda)\phi_b^{pb}(1 - \phi_b^{pb})(\bar{y} - \bar{x})}{\bar{x}(1 - \bar{x}) + \phi_b^{pb}(\bar{y} - \bar{x})[1 - 2\bar{x} - \phi_b^{pb}(\bar{y} - \bar{x})]^2} \geq 0$$

$$\frac{\partial[\phi_{b,N} - \phi_{b,E}]}{\partial(\bar{y} - \bar{x})} = \frac{(1 - \lambda)\phi_b^{pb}(1 - \phi_b^{pb})(\bar{z} - \bar{y})[\bar{x}(1 - \bar{x}) + \phi_b^{pb}(\bar{y} - \bar{x})^2]}{[\bar{x}(1 - \bar{x}) + \phi_b^{pb}(\bar{y} - \bar{x})[1 - 2\bar{x} - \phi_b^{pb}(\bar{y} - \bar{x})]^2]^2} \geq 0.$$

The two expressions are strictly positive when $\bar{z} > \bar{y} > \bar{x}$ and $\phi_b^{pb} \in (0, 1)$. ■

Lemma 6 *The difference between expected reputation after a failure in cases with and without scapegoating, $(\phi_{b,E} - \phi_b)$, is non decreasing in $(\bar{y} - \bar{x})$.*

Proof. For this proof, consider only the difference $\phi_{b,E} - \phi_b$ in equation (14) since, as shown in Lemma 5, $\frac{\partial[\phi_{b,N} - \phi_{b,E}]}{\partial(\bar{y} - \bar{x})} \geq 0$. Taking derivatives of $(\phi_{b,E} - \phi_b)$ with respect to $(\bar{y} - \bar{x})$,

$$\frac{\partial[\phi_{b,E} - \phi_b]}{\partial(\bar{y} - \bar{x})} = \frac{(1 - \lambda)\phi_b^{pb}(1 - \phi_b^{pb})^2(\bar{y} - \bar{x})[2\bar{x}(1 - \bar{x}) + (1 - 2\bar{x})\phi_b^{pb}(\bar{y} - \bar{x})]}{[\bar{x}[(1 - \bar{x}) + \phi_b^{pb}(\bar{y} - \bar{x})[1 - 2\bar{x} - \phi_b^{pb}(\bar{y} - \bar{x})]^2]^2} \geq 0,$$

which is non-negative because in the numerator, $(1 - \bar{x}) \geq (\bar{y} - \bar{x}) \geq \phi_b^{pb}(\bar{y} - \bar{x})$. This is also strictly positive whenever $\bar{z} > \bar{y} > \bar{x}$ and $\phi_b^{pb} \in (0, 1)$. ■

The difference in the blaming abilities between competents and inepts $(\bar{y} - \bar{x})$ basically measures the drop in expected reputation that, because of scapegoating, does *not* occur after a failure. Hence, an increase in $(\bar{y} - \bar{x})$ not only reduces the incentives to hire experts (by increasing $\phi_{b,E} - \phi_b$) but also increases the incentives to hire nonexperts (by increasing $\phi_{b,N} - \phi_{b,E}$).

Similarly, the difference in the abilities between blaming experts and nonexperts ($\bar{z} - \bar{y}$) measures the greater probability of having a positive court decision against employees from hiring nonexperts. Hence, an increase in $(\bar{z} - \bar{y})$ makes hiring nonexperts even more beneficial (by further increasing $\phi_{b,N} - \phi_{b,E}$).

3.4 Machiavellian Effect

The next proposition shows that, with scapegoating, efficient outcomes are more likely during good times rather than in bad times.

Proposition 7 Machiavellian Effect

Suppose $\lambda \in (0, 1)$, $\phi_0 \in [\lambda\theta, 1 - \lambda(1 - \theta)]$, $\delta \in (0, 1)$, and $\theta \in (0, 1)$, blaming reports are allowed, and competent have better blaming capabilities than inept ($\bar{y} > \bar{x}$). If there exist some $w > 0$ such that competent decide to hire experts in bad times, then they also decide to hire experts in good times, while the contrary is not true.

Proof. We need to show that $\Delta_G^S \geq \Delta_B^S$ by proving $X^G + \delta(1 - \lambda)V_f^G \geq X^B + \delta(1 - \lambda)V_f^B$ for all $\phi \in (0, 1)$. Considering equation (10), it suffices to show the following two claims:

Claim 1) $X^G \geq X^B$ for all ϕ .

Subtracting these expressions,

$$X^G - X^B = (1 - \alpha - \rho)[p(\phi_{b,N}) - p(\phi_{b,E})],$$

which is non negative, since $\alpha + \rho < 1$ by assumption; $p(\phi)$ is monotonic in ϕ and by equation (13) $\phi_{b,N} \geq \phi_{b,E}$.

Claim 2) $V_f^G \geq V_f^B$ for all ϕ .

Subtract these expressions:

$$V_f^G - V_f^B = \Pr(g|E)[Y_g^G - Y_g^B] + \Pr(b|E)[Y_b^G - Y_b^B],$$

is non negative, since $Y_i^G - Y_i^B = (1 - \alpha - \rho)(\bar{z} - \bar{y})[V(\phi_{bi}^{ec}) - V(\phi_{bi}^{sc})] \geq 0$ for $i \in \{g, b\}$. This is because $\alpha + \rho < 1$ and $\bar{z} \geq \bar{y}$ by assumption, $V(\phi)$ is monotonic in ϕ and $\phi_b^{ec} \geq \phi_b^{sc}$. Assuming scapegoating in the future as well does not change the conclusion.

Hence, in good times experts are hired for a wider range of wages w than in bad times. This does not imply a positive cutoff Δ_k^S , but it does imply it is more likely to have $\Delta_G^S > 0$ than $\Delta_B^S > 0$. In this way efficiency is more likely to be achieved in good times than in bad. ■

This Machiavellian effect arises because in bad times, superiors are more worried about potential reputation losses. Without scapegoating, this concern is irrelevant because, even when the probability of having a bad outcome is greater in bad times than in good, in both cases the differences in probabilities $(\alpha - \rho)$ are the same.²⁰ With scapegoating, however, hiring nonexperts become more attractive in bad times because they are easier to blame, buffering the reputation loss if a failure in fact occurs, exactly as proposed by Machiavelli.

4 Cheap Ways to Achieve an Efficient Equilibrium

We discussed how scapegoating generates inefficiencies by being nested after failures. A natural question arises. What happens with activities nested after successes? Many examples of this kind of situation exist in real life. In the sporting arena, All-Star Games, national teams and international championships (such as the Soccer World Cup) are organized for the participation of the best players. In organizations, corporations, and public offices, additional funds and responsibilities are assigned to divisions that outperform. In academic environments, round tables and plenary sessions at professional meetings are held by top researchers. In the entertainment industry, TV shows invite successful music and movie stars to exhibit their charisma or talent.

All these situations share the characteristic that individuals who are successful at their main activities gain access to additional stages that allow them to signal their competence even further. Even when society cares more about their main activities, events nested after successes may be very important to align reputation incentives more effectively, thus increasing the likelihood of efficient behavior.

4.1 Efficiency of Nested Activities after Successes

The model can be reinterpreted and modified to introduce nested activities after successes. Assume that instead of an irrelevant activity nested after a failure, such as scapegoating, the game is characterized by an irrelevant activity nested after successes. The structure of probabilities, timing, and parameters for the production stage have the same interpretation as before.

²⁰This is just an assumption to clarify the Machiavellian effect as much as possible. To assume otherwise does not change the main conclusion.

The difference appears in the second stage. After bad results the game ends, but after good results there is a nested activity, which can succeed (g, s) or fail (g, f) (these have basically the same spirit as (b, e_c) and (b, s_c) in the original model). Using similar notation as before, $\Pr(s|I, g, N) = x_g \in [0, \bar{x}_g]$, $\Pr(s|C, g, E) = y_g \in [0, \bar{y}_g]$, $\Pr(s|C, g, N) = z_g \in [0, \bar{z}_g]$. Differences in the abilities of being successful at the nested activity may exist.²¹

For example, if $\bar{y}_g > \bar{z}_g$, hiring experts increases the probability of being successful not only at producing but also at the additional nested activity. Superiors may choose the probabilities of being successful at the nested activity by eventually sabotaging workers' efforts. If there is no sabotage, the probability of success would be the maximum achievable (say, $y_g = \bar{y}_g$ if the employee is an expert), and a maximum sabotage intensity by the superior would eliminate the probability of success in the nested activity ($y_g = 0$). This decision about the sabotaging intensity has the same logic that decisions about blaming intensities have in the original model with scapegoating.

The equilibrium definition in this environment is the same as before, except that the three possible change in beliefs (in place of equations (5), (6), and (7)) are now ϕ_b (after a bad outcome), ϕ_g^s (after successes in both the first and second rounds), and ϕ_g^f (after success in production and failure in the nested unproductive activity). It is straightforward to show that in the efficient equilibrium, for all ϕ , after the first round $\phi_g > \phi > \phi_g$ and after the potential second round $\phi_g^s > (<)\phi_g > (<)\phi_g^f$ if $y_g > (<)x_g$. As in the scapegoating situation, only when $\bar{y}_g > \bar{x}_g$ may the nested stage generate a new reputation updating and affect efficiency.²² In what follows, unless stated otherwise, we consider only the relevant case $\bar{y}_g > \bar{x}_g$.

As in equations (13) and (14), it is also possible to define expected reputation after good results in the case of hiring experts and in the case of hiring nonexperts. Equations (13) and (14) could be restated as

$$\phi_{g,E} - \phi_{g,N} = \frac{(1 - \lambda)\phi_g^{pb}(1 - \phi_g^{pb})(\bar{y}_g - \bar{z}_g)(\bar{y}_g - \bar{x}_g)}{\bar{x}_g(1 - \bar{x}_g) + \phi_g^{pb}(\bar{y}_g - \bar{x}_g)[1 - 2\bar{x}_g - \phi_g^{pb}(\bar{y}_g - \bar{x}_g)]} > 0 \quad (17)$$

$$\phi_{g,E} - \phi_g = \frac{(1 - \lambda)\phi_g^{pb}(1 - \phi_g^{pb})^2(\bar{y}_g - \bar{x}_g)^2}{\bar{x}_g(1 - \bar{x}_g) + \phi_g^{pb}(\bar{y}_g - \bar{x}_g)[1 - 2\bar{x}_g - \phi_g^{pb}(\bar{y}_g - \bar{x}_g)]} > 0, \quad (18)$$

²¹Many of the examples discussed do not need delegation. In fact, in a better description competents would decide between exerting high or low efforts, while inepts would only be able to exert low efforts. This alternative environment, although the same in spirit, is different in that superiors would need to choose the effort level before both the first and the second rounds and pay twice the effort costs. Introducing this modification does not change the main conclusion.

²²If $\bar{y}_g \leq \bar{x}_g$, inepts prefer to sabotage the probability of success at the nested activity, imitating competents in order to be confused with them. In this way, inepts would not signal their own ineptitude (this is the same logic explained in Section 3.3).

where $\phi_g^{pb} = \Pr(C|g)$.

It is also informative to obtain the difference $\phi_{g,N} - \phi_g$

$$\phi_{g,N} - \phi_g = \frac{(1 - \lambda)(\phi_g^{pb})(1 - \phi_g^{pb})(\bar{y}_g - \bar{x}_g)[(\bar{z}_g - \bar{x}_g) - \phi_g^{pb}(\bar{y}_g - \bar{x}_g)]}{\bar{x}_g(1 - \bar{x}_g) + \phi_g^{pb}(\bar{y}_g - \bar{x}_g)[1 - 2\bar{x}_g - \phi_g^{pb}(\bar{y}_g - \bar{x}_g)]}, \quad (19)$$

which is positive only for $\phi_g^{pb} < \frac{\bar{z}_g - \bar{x}_g}{\bar{y}_g - \bar{x}_g}$.

A clear ordering exists among these expressions when $\bar{y}_g > \bar{x}_g$, as shown in Figure 3.

$$\phi_{g,E} > \phi_{g,N} > \phi_b \quad \text{and} \quad \phi_{g,E} > \phi_g > \phi_b \quad (20)$$

and, even when not relevant for our results,

$$\phi_{g,N} > \phi_g \quad \text{when} \quad \phi_g^{pb} < \frac{\bar{z}_g - \bar{x}_g}{\bar{y}_g - \bar{x}_g}.$$

Hiring experts increases expected reputation after good results ((18) is positive). Furthermore, hiring nonexperts also decreases expected reputation after good results ((19) is negative) for relatively high reputation levels. The final effect is always an increase in expected reputation gains from hiring experts.

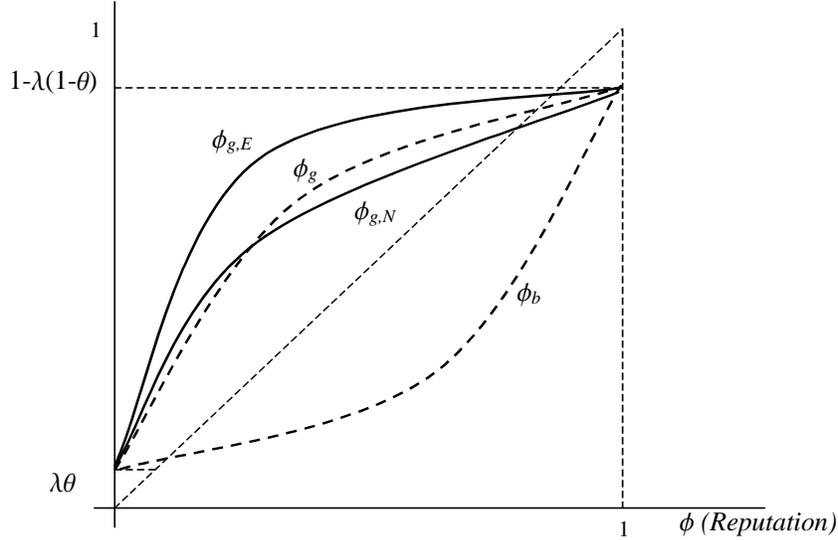
The main difference between nested activities after failures (such as scapegoating) and nested activities after successes is that, while the former reduces the expected reputation gains from good results (from $(\phi_g - \phi_b)$ to $(\phi_g - \phi_{b,E})$) and reduces the incentives to hire experts, the latter increases the expected reputation gains from good results (from $(\phi_g - \phi_b)$ to $(\phi_{g,E} - \phi_b)$) and increases the incentives to efficiently hire experts. Furthermore, while the former increases the incentives to hire nonexperts (by $(\phi_{b,N} - \phi_{b,E})$), the latter reduces them (by $(\phi_{g,E} - \phi_{g,N})$).

While Figure 3 delivers the basic intuition that sustains efficiency from nested activities after good results when $\bar{y}_g > \bar{x}_g$, formal proofs are very similar to the proofs for Propositions 2 and 3. Contrary to scapegoating, here it makes more sense to assume $\bar{y}_g > \bar{z}_g$. If this is the case, the irrelevant stage after successes always increases the likelihood of achieving an efficient equilibrium, without requiring a sufficient condition.

Summarizing, activities nested after successes align and exploit reputation concerns to achieve efficient outcomes, even when these activities are completely irrelevant and costless.²³

²³An extreme but illustrative example is the following. Suppose that hiring experts is efficient, but incentives from production are not enough for superiors to do it. Assume also that experts heavily outperform nonexperts at playing chess. A cheap way to achieve efficiency would be to introduce a chess game right after successes in production!

Figure 3: Expected reputation after the first round (with and without activities after successes)



Finally, the Machiavellian effects persists in this case. There are more incentives to hire experts in good times than in bad, since the probability of being successful and reaching the nested activity is greater in good times than in bad, further increasing the incentives to hire experts.

5 Conclusions

Scapegoating is a common behavior in public institutions, firms, sports, and even in the armed forces. The problem is not only the redistribution effects that unfair blaming may generate but also the possible introduction of inefficiencies in organizational performances. In fact, this conventional wisdom has been the main argument to promote some public reforms in the last decade, which were designed to assign more responsibility to superiors, thus reducing their chances to blame subordinates.

Inefficiencies may arise because of imperfect information, and reputation concerns may help to recover efficiency. However, scapegoating corrupts these attributes of reputation concerns by attenuating potential losses of reputation after failures, hence reducing the incentives to make costly decisions conducive to obtaining good results, such as hiring experts. Furthermore, scapegoating in fact increases the incentives to hire nonexperts in order to blame them more easily if something goes wrong.

To formalize this idea, I defined scapegoating as a non productive blaming activity "nested" after a bad result and introduced it as an extension of a reputation model. This "nested" activity may represent an additional way for superiors to signal their competence. Depending

on whether production or blaming is a better reputation builder, incentives to hire efficient workers will be affected. If blaming is a more secure way to build reputation, scapegoating reduces the incentives to hire experts, making the efficient situation more difficult to be sustained as an equilibrium.

Exploiting this nested reasoning, it may be better to locate activities after successes rather than after failures. If a principal only cares about the results in the first stage, it can be a good idea for him to introduce, right after positive results, activities in which competents outperform inepts. This will give superiors more incentives to achieve good results in the first stage in order to obtain the right to access the next stages and signal their competence even better. The model also delivers an interesting Machiavellian effect, for which competents tend to hire experts more in good times than in bad.

Even when this paper highlights a direct negative effect of scapegoating, there are potential arguments in its favor. This model focuses on only one particular delegation motive, which is reputation, and leaves out other important reasons such as specialization and scale. This is why conclusions are biased toward the assignment of complete responsibility to superiors. A more comprehensive model, considering all determinants, would be necessary to obtain the optimal allocation of responsibility and accountability to superiors.

These caveats aside, the model supports the general view that blaming possibilities after failures reduces the incentives to make efficient decisions that reputation concerns introduce. The model also proposes ways to reinforce reputational incentives by the use of, eventually irrelevant but cheap, activities nested after certain results.

References

- Alesina, Alberto, and Guido Tabellini. 2005. "Why Do Politicians Delegate?" Working Paper 11531. National Bureau of Economic Research.
- . 2007. "Bureaucrats or Politicians? Part I: A Single Policy Task." *American Economic Review* 97:169–179.
- . 2008. "Bureaucrats or Politicians? Part II: Multiple Policy Tasks." *Journal of Public Economics* 92:425–447.
- Bell, Nancy E., and Philip E. Tetlock. 1989. "The Intuitive Politician and the Assignment of Blame in Organizations." In *Impression Management in the Organization*, edited by Robert Giacalone and Paul Rosenfel, 105–123. Hillsdale, NJ: Erlbaum.
- Cole, Hal, and Patrick Kehoe. 1996. "Reputation Spillover Across Relationships: Reviv-

- ing Reputation Models of Debt." Working Paper 5486, National Bureau of Economic Research.
- Cripps, Martin, George Mailath, and Larry Samuelson. 2004. "Imperfect Monitoring and Impermanent Reputations." *Econometrica* 72:407–432.
- . 2007. "Disappearing Private Reputations in Long-Run Relationships." *Journal of Economic Theory* 134:287–316.
- Dezso, Cristian. 2004. "Scapegoating and Firm Reputation." Mimeo, Stern School of Business, New York University.
- Douglas, Tom. 1995. *Scapegoating: Transferring Blame*. London: Routledge.
- Fudenberg, Drew, and David K. Levine. 1989. "Reputation and Equilibrium Selection in Games with a Patient Player." *Econometrica* 57:759–778.
- . 1992. "Maintaining a Reputation When Strategies Are Imperfectly Observed." *Review of Economic Studies* 59:561–579.
- Huson, Mark R., Paul H. Malatesta, and Robert Parrino. 2004. "Managerial Succession and Firm Performance." *Journal of Financial Economics* 74:237–275.
- Kreps, David, and Robert Wilson. 1982. "Reputation and Imperfect Information." *Journal of Economic Theory* 27:253–279.
- Mailath, George, and Larry Samuelson. 1998. "Your Reputation Is Who You're Not, Not Who You Would Like to Be." CARESS, Working Paper 98-11, University of Pennsylvania.
- . 2001. "Who Wants a Good Reputation?" *Review of Economic Studies* 68:415–441.
- . 2006. *Repeated Games and Reputations*. New York: Oxford University Press.
- Martin, John. 1997. "Changing Accountability Relations: Politics, Consumers and the Market." Public Management Service, Organisation for Economic Cooperation and Development.
- Milgrom, Paul, and John Roberts. 1982. "Predation, Reputation and Entry Deterrence." *Journal of Economic Theory* 27:280–312.
- Polidano, Charles. 1999. "The Bureaucrat Who Fell Under a Bus: Ministerial Responsibility and the Derek Lewis Affair in Britain." *Governance: An International Journal of Policy, Administration and Institutions*. 12:201–229.
- Segendorff, Bjorn. 2000. "A Signalling Theory of Scapegoats." Mimeo, Department of Economics, Stockholm School of Economics.
- Winter, Eyal. 2001. "Scapegoats and Optimal Allocation of Responsibilities." Mimeo, Center for Rationality and Interactive Decision Theory, The Hebrew University of Jerusalem.

A Appendix

Proof for Proposition 2. Fix ϕ and suppose an efficient situation (i.e., competent superiors always choose to hire experts in both good times and bad ($\tau_G(\phi) = \tau_B(\phi) = 1$)). It is straightforward to show that for all ϕ , after the first round $\Pr(C|g) > \Pr(C) > \Pr(C|b)$.

Hence, given any state $k \in \{B, G\}$, for all feasible ϕ , $\varphi(\varphi(\phi|g)|g) = \phi_{gg} > \phi_g > \phi > \phi_b > \phi_{bb}$ and $\phi_{gi} > \phi_{bi}$ for $i \in \{g, b\}$.

If $k = B$, competents' value function when hiring experts is

$$V(\phi, E) = p(\phi) - w + \delta(1 - \lambda)[\Pr(g|E, B)V(\phi_g) + \Pr(b|E, B)V(\phi_b)]$$

$$V(\phi, E) = p(\phi) - w + \delta(1 - \lambda)[\alpha V(\phi_g) + (1 - \alpha)V(\phi_b)].$$

The payoff from deviating by hiring a nonexpert and thereafter playing the equilibrium strategy of hiring experts is

$$V(\phi; N) = p(\phi) + \delta(1 - \lambda)[\rho V(\phi_g) + (1 - \rho)V(\phi_b)].$$

Thus,

$$V(\phi, E) - V(\phi; N) = -w + \delta(1 - \lambda)[X] + \delta^2(1 - \lambda)^2\{\Pr(g|E)Y_g + \Pr(b|E)Y_b\},$$

where

$$X = (\alpha - \rho)[p(\phi_g) - p(\phi_b)]$$

$$Y_i = (\alpha - \rho)[V(\phi_{gi}) - V(\phi_{bi})] \quad \text{for } i \in \{g, b\}$$

$$\Pr(g|E) = \alpha + \gamma(1 - \rho - \alpha) = 1 - \Pr(b|E).$$

In order for $V(\phi, E) - V(\phi; N) \geq 0$, it is necessary that

$$w \leq \delta(1 - \lambda)[X + \delta(1 - \lambda)V_f]; \quad \text{for all } \phi \in [\lambda\theta, 1 - \lambda(1 - \theta)],$$

where $V_f = \Pr(g|E)Y_g + \Pr(b|E)Y_b$.

Then we can define Δ_B^{NS} as the minimum value of the expression $\delta(1 - \lambda)[X + \delta(1 - \lambda)V_f]$ over the range $\phi \in [\lambda\theta, 1 - \lambda(1 - \theta)]$

$$\Delta_B^{NS} = \min_{\phi \in [\lambda\theta, 1 - \lambda(1 - \theta)]} \left\{ \delta(1 - \lambda)[X + \delta(1 - \lambda)V_f] \right\}. \quad (21)$$

If $k = G$, the condition for competents to hire experts is the same as (21). This is because ϕ_g and ϕ_b do not change and $(\alpha - \rho)$ is by assumption the same. Then, it is always possible to find some $\Delta^{NS} \geq w > 0$ such that $\Delta^{NS} = \Delta_G^{NS} = \Delta_B^{NS}$ and competents hire experts for all ϕ .

Finally, I show that $\delta(1 - \lambda)[X + \delta(1 - \lambda)V_f]$ is positive for all $\phi \in (0, 1)$, such that $\Delta^{NS} > 0$.

- $\delta(1 - \lambda) > 0$, since $\delta > 0$ and $\lambda < 1$.
- $X = (\alpha - \rho)[p(\phi_g) - p(\phi_b)] > 0$, since $\alpha > \rho$ and $p(\phi)$ is monotonically increasing in ϕ ($\frac{\partial p(\phi)}{\partial \phi} = \alpha - \rho > 0$).²⁴
- $V_f > 0$, since $Y_g > 0$ and $Y_b > 0$ (the value function V is monotonically increasing in ϕ).²⁵ Hence, $\Delta^{NS} > 0$ for all $\phi \in [\lambda\theta, 1 - \lambda(1 - \theta)]$.

■

Proof for Proposition 3. This proof proceeds in two steps. First, I solve for superiors' optimal blaming intensities ($x(\phi)$, $y(\phi)$, and $z(\phi)$) consistent with consumers' beliefs. Second, using these results from the blaming stage, we derive conditions for an efficient equilibrium.

Step 1: Blaming stage equilibrium

First, consider as given consumers' beliefs about blaming intensities and determine optimal decisions by superiors (both competents (y) and inepts (x)).²⁶ Second, considering the optimal blaming intensities, we check if beliefs are correct and consistent with those strategies.

Blaming decisions are made by superiors knowing their own type, their previous delegation choices, and the state of nature. For example, in bad times, when competents hired experts and decided a blaming intensity y ,

$$V(\phi, y) = p(\phi) - w + \delta(1 - \lambda)[\alpha V(\phi_g) + (1 - \alpha)[yV(\phi_b^{ec}) + (1 - y)V(\phi_b^{sc})]],$$

where ϕ_b^{ec} and ϕ_b^{sc} are given by consumers' beliefs about y and x .

For any deviation from y , say to y' , we can define

$$VD(y) = V(\phi, y') - V(\phi, y) = \delta(1 - \lambda)(1 - \alpha)(y' - y)[V(\phi_b^{ec}) - V(\phi_b^{sc})].$$

1) Assume consumers believe $y = x$. By equations (6) and (7), $\phi_b^{ec} = \phi_b^{sc} = \phi_b$. Since $V(\phi_b^{ec}) - V(\phi_b^{sc}) = 0$, competents are indifferent between choosing any $y' \in [0, \bar{y}]$ (regardless of $(y' - y)$, always $VD(y) = 0$). Similarly, inepts are indifferent between choosing any $x' \in [0, \bar{x}]$ (regardless of $(x' - x)$, $VD(x) = 0$). Hence, consumers' beliefs $y = x$ are correct and consistent with equilibrium strategies, supporting multiple pooling equilibria in which no further reputation update is obtained from the blaming activity.

²⁴More specifically, as in Mailath and Samuelson (2001), suppose F and G are two distributions describing consumers' beliefs over the delegation decisions by competents in period t . If F first-order stochastically dominates G , then superiors' revenues in period t under F are higher than they are under G .

²⁵Following Mailath and Samuelson (2001), let $f_t(\phi, \phi_0, t_0)$ be the distribution of consumers' posteriors ϕ at time $t > t_0$ induced by strategy τ given period- t_0 posteriors ϕ_0 . Then, $f_t(\phi, \phi_0, t_0)$ first-order stochastically dominates $f_t(\phi, \phi'_0, t_0)$ for all $t > t_0$ and $\phi_0 > \phi'_0$. The same idea is true for the distribution of revenues. Hence, $V(\phi)$ is monotonic.

²⁶Recall at this point that z is not relevant for consumers to update beliefs, since we are focusing only on efficient equilibria in which competents always hire experts.

2) Assume consumers believe $y > x$. Then, $\phi_b^{ec} > \phi_b > \phi_b^{sc}$. Since $V(\phi_b^{ec}) - V(\phi_b^{sc}) > 0$, competents choose $y' = \bar{y}$, which maximizes $VD(y)$. Similarly, inepts will choose $x' = \bar{x}$. Only consumers' beliefs $y = \bar{y}$ and $x = \bar{x}$ will be correct, which are consistent with beliefs $y > x$ solely when $\bar{y} > \bar{x}$ is the only separating equilibrium in which the blaming activity represents an additional reputation updating.

3) Assume consumers believe $y < x$. Then, $\phi_b^{ec} < \phi_b < \phi_b^{sc}$. Since $V(\phi_b^{ec}) - V(\phi_b^{sc}) < 0$, competents choose $y' = 0$ and inepts $x' = 0$. Only consumers' beliefs $y = 0$ and $x = 0$ will be correct, which is not consistent with beliefs in which $y < x$. This case cannot be an equilibrium.²⁷

Step 2: Delegation stage equilibrium

1) Let $\bar{y} \leq \bar{x}$.

Fix ϕ and suppose an efficient situation ($\tau_G(\phi) = \tau_B(\phi) = 1$). Since the only possible equilibrium in the blaming stage is a pooling one, where $\phi_b^{ec} = \phi_b^{sc} = \phi_b$, we have exactly the same expressions used to obtain equilibrium conditions without scapegoating (in the proof for Proposition 2'). Hence, under $\bar{y} \leq \bar{x}$, scapegoating does not affect efficiency conditions.

2) Let $\bar{y} > \bar{x}$.

Even when pooling equilibria in blaming intensities that do not affect efficiency conditions exist in this situation, we will focus on the unique separating equilibrium in which $y = \bar{y}$, $x = \bar{x}$, and $z = \bar{z}$ such that $\phi_b^{ec} > \phi_b > \phi_b^{sc}$.

Fix ϕ and suppose an efficient situation ($\tau_G(\phi) = \tau_B(\phi) = 1$). It is straightforward to show that for all ϕ , after the first round $\Pr(C|g) > \Pr(C) > \Pr(C|b)$ and after the potential second round $\Pr(C|b, e_c) > (<) \Pr(C|g) > (<) \Pr(C|b, s_c)$ if $y > (<) x$. Given any state $k \in \{B, G\}$, for all feasible ϕ , $\varphi(\phi|g|g) = \phi_{gg} > \phi_g > \phi > \phi_b > \phi_{bb}$ and $\phi_{gi} > \phi_{b_1i} > \phi_{bi} > \phi_{b_2i}$ for $i \in \{g, b_1, b_2\}$, with $b_1 = (b, e_c)$ and $b_2 = (b, s_c)$.

If $k = B$, competents' value function when hiring experts is

$$V(\phi, E) = p(\phi) - w + \delta(1 - \lambda)[\alpha V(\phi_g) + (1 - \alpha)[\bar{y}V(\phi_b^{ec}) + (1 - \bar{y})V(\phi_b^{sc})]],$$

and from deviation by delegating to a nonexpert is

$$V(\phi, N) = p(\phi) + \delta(1 - \lambda)[\rho V(\phi_g) + (1 - \rho)[\bar{z}V(\phi_b^{ec}) + (1 - \bar{z})V(\phi_b^{sc})]].$$

Thus,

$$\begin{aligned} V(\phi, E) - V(\phi; N) = & -w + \delta(1 - \lambda)[X^B] \\ & + \delta^2(1 - \lambda)^2 \{ \Pr(g|E)Y_g^B + \Pr(b|E)[\bar{y}Y_{b_1}^B + (1 - \bar{y})Y_{b_2}^B] \}, \end{aligned}$$

²⁷Because we are focusing on efficient equilibria, we checked beliefs for x and y , but a competent type that deviated in the first stage hiring nonexperts will also choose any $z \in [0, \bar{z}]$ in 1), $z = \bar{z}$ in 2), and $z = 0$ in 3).

where

$$X^B = (\alpha - \rho)p(\phi_g) + (1 - \alpha)[\bar{y}p(\phi_b^{ec}) + (1 - \bar{y})p(\phi_b^{sc})] \\ - (1 - \rho)[\bar{z}p(\phi_b^{ec}) + (1 - \bar{z})p(\phi_b^{sc})]$$

$$Y_i^B = (\alpha - \rho)V(\phi_{gi}) + (1 - \alpha)[\bar{y}V(\phi_{bi}^{ec}) + (1 - \bar{y})V(\phi_{bi}^{sc})] \\ - (1 - \rho)[\bar{z}V(\phi_{bi}^{ec}) + (1 - \bar{z})V(\phi_{bi}^{sc})]$$

for $i \in \{g, b_1, b_2\}$

and, as in the previous proof, $\Pr(g|E) = \alpha + \gamma(1 - \rho - \alpha) = 1 - \Pr(b|E)$.

At this point, it is useful to express X^B in terms of expected reputation after the hiring decision.

Let's define

$$\phi_{b,E} = \bar{y}\phi_b^{ec} + (1 - \bar{y})\phi_b^{sc} \quad (22)$$

$$\phi_{b,N} = \bar{z}\phi_b^{ec} + (1 - \bar{z})\phi_b^{sc}. \quad (23)$$

Hence,²⁸

$$X^B = (\alpha - \rho)p(\phi_g) + (1 - \alpha)p(\phi_{b,E}) - (1 - \rho)p(\phi_{b,N})$$

An equilibrium in which competents only hire experts when $k = B$ requires that $V(\phi, E) - V(\phi; N) \geq 0$ for all feasible reputation measures ϕ . A necessary condition is that cost differences w fulfill

$$w \leq \delta(1 - \lambda)[X^B + \delta(1 - \lambda)V_f^B]; \text{ for all } \phi \in [\lambda\theta, 1 - \lambda(1 - \theta)],$$

where $V_f^B = \Pr(g|E)Y_g^B + \Pr(b|E)[\bar{y}Y_{b_1}^B + (1 - \bar{y})Y_{b_2}^B]$.

Then we can define Δ_B^S as the minimum value of the expression $\delta(1 - \lambda)[X^B + \delta(1 - \lambda)V_f^B]$ over the range $\phi \in [\lambda\theta, 1 - \lambda(1 - \theta)]$

$$\Delta_B^S = \min_{\phi \in [\lambda\theta, 1 - \lambda(1 - \theta)]} \left\{ \delta(1 - \lambda)[X^B + \delta(1 - \lambda)V_f^B] \right\} \quad (24)$$

To save notation, it is possible to assume a case in which the future does not present scapegoating possibilities, so there is just one current shot at blaming. In this case, from tomorrow on, it would be possible to have only two possible states $i \in \{g, b\}$. It is straightforward to check that (24) is simplified to $V_f^B = \Pr(g|E)Y_g^B + \Pr(b|E)Y_b^B$. This last expression is used in Proposition 3.

If $k = G$, the proof is identical to the previous one but having $\Pr(g|E, G) = (1 - \rho)$ and $\Pr(g|N, G) = (1 - \alpha)$ instead.

²⁸It is not possible to do the same for Y_i^B because we do not know the form of the value functions, just their monotonicity in ϕ (recall we are not assuming linearity of $V(\phi)$).

Then,

$$X^G = (\alpha - \rho)p(\phi_g) + \rho p(\phi_{b,E}) - \alpha p(\phi_{b,N})$$

$$Y_i^G = (\alpha - \rho)V(\phi_{gi}) + \rho[\bar{y}V(\phi_{bi}^{ec}) + (1 - \bar{y})V(\phi_{bi}^{sc})] - \alpha[\bar{z}V(\phi_{bi}^{ec}) + (1 - \bar{z})V(\phi_{bi}^{sc})]$$

for $i \in \{g, b_1, b_2\}$.

Hence, the condition for competents to hire experts and to achieve the efficient outcome as an equilibrium is

$$w \leq \delta(1 - \lambda)[X^G + \delta(1 - \lambda)V_f^G]; \text{ for all } \phi \in [\lambda\theta, 1 - \lambda(1 - \theta)],$$

where $V_f^G = \Pr(g|E)Y_g^G + \Pr(b|E)[\bar{y}Y_{b_1}^G + (1 - \bar{y})Y_{b_2}^G]$.

Then we can define Δ_G^S as the minimum value of the expression $\delta(1 - \lambda)[X^G + \delta(1 - \lambda)V_f^G]$ over the range $\phi \in [\lambda\theta, 1 - \lambda(1 - \theta)]$

$$\Delta_G^S = \min_{\phi \in [\lambda\theta, 1 - \lambda(1 - \theta)]} \left\{ \delta(1 - \lambda)[X^G + \delta(1 - \lambda)V_f^G] \right\}. \quad (25)$$

As in the previous case, to save notation $V_f^G = \Pr(g|E)Y_g^G + \Pr(b|E)Y_b^G$ in equation (25), which is used in Proposition 3. ■

Extension of proof for Proposition 4. Considering blaming activities in current and future periods, the only difference arises in the definition of V_f^G , after equation (25).

Hence, we need to prove $V_f - V_f^G \geq 0$, where

$$V_f^G = \Pr(g|E)Y_g^G + \Pr(b|E)[\bar{y}Y_{b_1}^G + (1 - \bar{y})Y_{b_2}^G].$$

$$\text{Then, } V_f - V_f^G = \Pr(g|E)[Y_g - Y_g^G] + \Pr(b|E)[Y_b - \bar{y}Y_{b_1}^G - (1 - \bar{y})Y_{b_2}^G].$$

This expression will be non-negative whenever $Y_g - Y_g^G \geq 0$ (proved in Proposition 3) and $Y_b - \bar{y}Y_{b_1}^G - (1 - \bar{y})Y_{b_2}^G \geq 0$, which happens under two sufficient conditions,

1) $(\alpha - \rho)[\bar{y}V(\phi_{gb}) + (1 - \bar{y})V(\phi_{gb})] \geq (\alpha - \rho)[\bar{y}V(\phi_{gb_1}) + (1 - \bar{y})V(\phi_{gb_2})]$, or

$$\bar{y} \leq \frac{V(\phi_{gb}) - V(\phi_{gb_2})}{V(\phi_{gb_1}) - V(\phi_{gb_2})} \leq 1,$$

since $\phi_{gb_1} \geq \phi_{gb} \geq \phi_{gb_2}$.

2) $(\alpha\bar{z} - \rho\bar{y})[\bar{y}V(\phi_{bb_1}^{ec}) + (1 - \bar{y})V(\phi_{bb_2}^{ec})] \geq (\alpha - \rho)V(\phi_{bb})$

The sufficient conditions for this to hold are both (as in (15))

$$\bar{y} \geq \frac{V(\phi_{bb}) - V(\phi_{bb_2}^{ec})}{V(\phi_{bb_1}^{ec}) - V(\phi_{bb_2}^{ec})} \text{ and } \bar{z} > z^* = 1 - \frac{\rho}{\alpha}(1 - \bar{y}).$$

■