Don't ask why things went wrong

Nested Reputation and Scapegoating Inefficiency†

Guillermo L. Ordoñez @

UCLA - September 2005

Abstract

Scapegoating is a source of inefficiency in organizations. This is a commonly held view. But despite its commonality and its use as a justification for reforms designed to reduce scapegoating, no formal model has been developed to explain why and when this is the case.

We propose an explanation by focusing on how superiors make decisions inside an organization when they only care about reputation. Consider delegation choices, for example. The hiring of 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 since they are easier to blame if something goes wrong.

In this context, scapegoating is a "nested" activity that only occurs after bad outcomes. Its results do not directly affect the welfare of society but indirectly affect the decisions governing the probability of success in production.

We also examine "nested" activities following good outcomes and show how these can increase efficiency without relying on costly incentives.

This "nested reputation" game also predicts that in good times superiors tend to hire more efficient workers than in bad times, a feature we call "Machiavellian Effect".

Keywords: scapegoating, reputation, delegation, efficiency.
JEL Codes: D81, D21, D73, M51

[0]† Many thanks to David K. Levine and Vasiliki Skreta for excellent suggestions and to participants at UCLA Economic Theory Proseminar for comments. Potential errors are mine.

[0]@ UCLA, Economics Department. guilord@ucla.edu

1 Introduction

A quick review of newspapers in many countries shows that people condemn scapegoating behavior. This attitude is not only due to the unfairness scapegoating represents, but also to its negative effects on efficiency and performance in organizations.

It is typically believed that a superior's impunity to blame subordinates plays a big role in the explanation of inefficiencies in both public and private sectors. The common argument is that scapegoating leads to poor performance because superiors think *"Why exert high efforts if I can always blame an employee in case things go wrong?"*.

In fact, this general idea has been widely mentioned in recent reforms designed to change accountability relations in the public sector. For example, the assignment of more responsibility to superiors has been a main objective of reforms made by OECD countries over the past decade. "Better accountability, it is often suggested, is both an end in itself (representing democratic values) and a means towards the development of more efficient and effective organizations" (Martin, 1997).

Specific cases are the "Next Steps" and the "Outcome-Output" programs of the UK and New Zealand respectively, where the goal is to generate "(mechanisms of) defense available to bureaucrats in danger of becoming political scapegoats" (Polidano, 1999).

Another criticism to scapegoating in public service comes from the Rome Conference for an International Criminal Court. It was stated, for example, that "(given a subordinate crime), civilian superiors are not subjected to the "should have known" standard of their military counterparts"[1].

But even when these movements have proposed the reduction of superior's scapegoating to improve efficiency, no model has been developed so far to formalize this seemingly well-accepted idea. The unique, but yet informal, explanation offered to justify the reforms is that superiors exert less effort when covered by their employees if something fails. But this explanation is criticized by arguing that under scapegoating employees may also want to increase efforts, making the final results unclear.

To understand the consequences of scapegoating on efficiency, we abstracts any effort consideration from this paper and focus only on delegation decisions made by reputation-concerned superiors.

The two forces almost exclusively used to explain delegation decisions are specialization (because of information and skills restrictions to superiors) and scale (because of the impossibility of superiors to effectively oversee all different parts of the production process). But another consideration, almost always forgotten by the economic literature, is the reputation concerns of those who make the decisions (Bendor et al. (2001)).

[1] Article 25 of the Responsibility of Commanders and Superiors Statute

In fact, reputation motifs of delegation were carefully discussed a long time ago by Machiavelli in his famous book "The Prince". One of his dictums says, *"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, maintaining their reputation and the love of the kingdom's people.

More recently, Alesina and Tabellini (2005) formally modeled this pattern. They show that politicians tend to retain under their control policy tools that are useful for building winning coalitions or generating campaign contributions. Contrarily, they tend to delegate tasks that expose them to risk, in order to use bureaucrats as scapegoats or "risk shields" if something goes wrong. The authors use as examples European national politicians who publicly blame bureaucrats in front of the European Commission and big bureaucracies (such as the IMF) that sometimes are used as international scapegoats.[2]

With these exceptions in mind, delegation and reputation models have generally been divorced in the literature, which is surprising considering the existence of outstanding studies that provide formal models for conceptualizing and understanding reputation rigorously. Starting with Kreps and Wilson (1982) and Milgrom and Roberts (1982), important contributions were Fudenberg and Levine (1989 and 1992), Mailath and Samuelson (2001) and Cripps, Mailath and Samuelson (2004).

To forget reputation when explaining delegation is not a trivial omission since, depending on how reputation is built, it may be corrosive to optimal delegation choices. For example, if a good reputation is not only obtained from being successful at producing but also from being successful at blaming, superiors may prefer to hire inefficient workers who are easier to blame rather than hire efficient workers who are better to produce. In this case, scapegoating can be used strategically by superiors who care about reputation, affecting delegation decisions and efficiency results.

Despite the recognition of strategic reasons for scapegoating in the socialpsychology literature, (Bell & Tetlock, 1989; Douglas, 1995), formal economic studies of this behavior are new and sparse. Dezso (2004) analyzes the conditions under which random firing of potential innocents (scapegoats) is a reaction to failures in order to maintain reputation. He focuses in firing and not in hiring, without being able to analyze the impact on efficiency. Segendorff (2000) analyzes the possible hiring of scapegoats using a signalling game, without analyzing efficiency consequences either. 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 did not consider reputation effects nor hiring decisions though.

Empirical studies about scapegoating are even less common. An exception is

[2] See also Alesina and Tabellini (2004)

Huson et al. (2004) who developed a moral hazard-driven scapegoat hypothesis based on agency models to study the impact of managerial succession on firm performance.

In this paper, we offer a novel interpretation of scapegoating as a non-productive activity that occurs only after failures and may be used by reputation-concerned superiors to signal their competence. To work with this idea we extend the reputation environment developed by Mailath and Samuelson (2001)[3] introducing both delegation and scapegoating.

Superiors will be either competents or inepts. While competents have the possibility of hiring both efficient (expert) and inefficient (nonexpert) workers, inepts can only work with inefficient ones. When deciding whether or not to hire an expert, competents compare the higher probability of obtaining good results and gaining reputation to the higher wages they have to pay

In the same way that production outcomes are useful elements for inferring a superior's competence, scapegoating is introduced as an additional alternative superiors can use to signal their capability. If blaming is a clearer way to signal competence than production, superiors will prefer to make decisions that exploit blaming, not caring if only production matters to society. In this way superiors will make delegation decisions not only thinking on production but also on potential scapegoating.

Hence, conditions for an efficient equilibrium with and without scapegoating are compared, concluding that in the former case it is more difficult, or even impossible, to achieve efficiency as an equilibrium.

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 reputation gains from working with efficient employees. 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 inefficient employees.

But in order scapegoating really affects hiring decisions, competents must have both better production possibilities and better blaming capabilities. If this is the case, they will prefer to hire nonexperts in situations where, without the opportunity to abdicate responsibility, would have preferred to hire experts.

Even when focusing on how scapegoating, an activity after failures, affects efficiency, we also extend the logic to study activities that are better at building reputation as well, but that only occur after successes. Many examples, from areas so diverse as sports and universities, are discussed in the paper.

In fact, we will propose to increase the possibilities for achieving efficiency by introducing unproductive stages after successes in nested reputation games, exploiting reputation forces in the right direction without requiring monetary resources or costly incentives.

[3] See also this paper for a nice discussion about similarities and differences with more standard reputation models from Fudenberg and Levine (1989, 1992).

Finally, we will show also that considering nested reputation games it is possible to recover a "Machiavellian Effect". Our version of his dictum will be *"Superiors tend to hire nonexperts in bad times and experts in good times"*.

To my knowledge, nested reputation models do not exist, constituting this work an initial effort to understand how the results in a standard reputation game change when introducing "nested" activities. Even when this paper shares some features with the literature on "reputation spillover," the logic is not the same[4]. Spillover refers to the impact of reputation in one aspect on the reputation in another aspect, hence dealing with different and multiple types of reputation. Here reputation is based on a single aspect but it is constructed through several nested stages and steps.

The next section presents the basic model of reputation with delegation and scapegoating, being an special application of a "nested reputation" environment. Section 3 analyzes the conditions for an efficient equilibrium, showing how and when scapegoating leads to inefficiency and to "Machiavellian Effects". Section 4 proposes a way to induce efficiency by exploiting reputation concerns, without relying on the use of costly incentives. Section 5 concludes.

2 The Model

2.1 Description

This model extends Mailath and Samuelson (2001) by introducing scapegoating as a nested activity and delegation.

Assume a superior (who can be a country president, 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 (DE), paying a wage w_E, or to hire nonexperts (DN), paying $w_N < w_E$. Inepts can only delegate to nonexperts because no expert would like to work with them. Before deciding, the superior observes the state of the nature, good (G) or bad (B), which affects the probability of success in achieving the target.

After the decision is made, production happens and a non-deterministic output, which can be good (g) or bad (b), is obtained. When competents hire experts the probability of a good result in good times is $(1-\rho) > \frac{1}{2}$ and in bad times $\alpha < \frac{1}{2}$. Hiring nonexperts allows them to obtain good results with a probability $(1-\alpha)$ in good times and ρ in bad times (where $\alpha > \rho$)[5]. An

[4] See a discussion on Cole and Kehoe (1996).

[5] The assumption of symmetry in probabilities does not change the main conclusions and allows the use of just two parameters (α and ρ) instead of four, eliminating awkward expressions.

important assumption is that $(\alpha - \rho) > (w_E - w_N) > 0$, which means it is efficient for society that competents always hire experts, both in good and bad times. It basically says that, if "consumers" knew agents' delegation decisions, they would be willing to pay a premium for competents to always hire experts.

When the outcome is finally observed and the result is a failure, the superior has to make a report about its causes, deciding 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 and not after successes. Once the report is done, a non-deterministic 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) is considered the culprit of the failure.

Deciding the intensity of the blaming and the amount of evidence displayed, superiors choose directly the probability of the "court" blaming the worker[6]. For example, inepts decide a probability y the "court" pronounces against subordinates such that $y \in [0, \overline{y}]$ where $\overline{y} \leq 1$. This means there can be a maximum capacity to successfully blame workers, or which is the same, maximum blaming intensities do not necessarily assure the "court" deciding against subordinates.

When competents choose a probability the "court" decides in his favor, they know if the blamed employee is an expert or a nonexpert. If competents worked with experts, they decide a probability x and if they worked with nonexperts they may choose a different one, z. These probabilities will be such that $x \in [0, \overline{x}]$ and $z \in [0, \overline{z}]$ where $\overline{x} \leq \overline{z} \leq 1$.

The maximum probabilities of successful blaming in all cases ($\overline{x}$, $\overline{y}$ and $\overline{z}$) are exogenous parameters known by everybody in the economy and basically describe blaming capabilities under maximum blaming intensities.

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 $\theta \in (0, 1)$ [7].

"Consumers" (continuum of identical persons of unit mass such that no single individual can affect the future play of the game) repeatedly receive the output generated under superior's commands (e.g, consumers purchase a good, citizens receive a service and stockholders obtain dividends). 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 results.

Even when "consumers" know the probability of being in a good state is $\Pr(G) = \gamma$, they are not able to see if the economy is in good or bad times nor if the superior hired experts or nonexperts. "Consumers" can only see the results from production activities (success or failure) and from blaming activities after failures (superior or employee considered culprit).

[6] A nil 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 the "court" pronounces against employees.

[7] This assumption is needed to sustain an efficient equilibrium in the long run, as discussed in Mailath and Samuelson (2001) and Cripps, Mailath and Samuelson (2004).

From this information they update the probability that the superior is competent, $\Pr(C) = \phi$, (i.e. his or her reputation). This is of the utmost importance to superiors since we assume 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), 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 the reason superiors are so concerned about reputation while "consumers" are only concerned about the utility derived from production.

2.2 Timing

The timing of the model is:

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

2) The superior observes ϕ, w_E, w_N and the environment state (G or B). Competents decide between hire experts or nonexperts. Inepts can only hire nonexperts. "Consumers" do not observe this decision, nor whether there are good or bad times.

3) 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 realization of utility outcome, which is public.

4) 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. 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

In the presence of uncertainty about superior's type, the state variable is just the probability assigned by "consumers" to the superior being competent (i.e. the reputation denoted as ϕ). A Markov strategy for competents in good times (G) is a mapping $\tau_G : [0, 1] \rightarrow [0, 1]$, where $\tau_G(\phi)$ is the probability of hiring an expert when reputation is $\phi \in [0, 1]$. Similarly, in bad times (B) the Markov strategy for competents is a mapping $\tau_B : [0, 1] \rightarrow [0, 1]$. Inepts make no choice, having then a trivial strategy of hiring nonexperts.[8]

[8] As noted in Mailath and Samuelson (2001), by restricting attention to strategies that only depend on consumers' posteriors, in equilibrium different superiors will behave identically in identical situations.

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

In a Markov perfect equilibrium superiors maximize revenues, "consumers"' expectations are correct and "consumers" use a Bayes' rule to update their posterior probabilities.

Since the state variable is the reputation ϕ, the model relies importantly on the updating of beliefs about the competence of the superior. There are two rounds of updating that follow a Bayes rule: The update after production ($\Pr(C|g)$ and $\Pr(C|b)$) and the potential update ONLY after a bad outcome ($\Pr(C|b, e_c)$ and $\Pr(C|b, s_c)$), which is based on the observation of "court"'s decisions after scapegoating. The following lemmas show explicitly in terms of parameters and decision rules the updating after the two possible states in each round. Proofs are in the Appendix.

Lemma 1 ***Beliefs updating after production (first round)***

After production (first stage), two updates can occur

a) The production is successful and the outcome is good (g)

$$\Pr(C|g) = \frac{\Pr(g|C)\phi}{\Pr(g|C)\phi + \Pr(g|I)(1-\phi)} \tag{1}$$

where

$$\Pr(g|C) = \gamma[(1-\rho)\tau_G + (1-\alpha)(1-\tau_G)] + (1-\gamma)[\alpha\tau_B + \rho(1-\tau_B)]$$

$$\Pr(g|I) = \gamma(1-\alpha) + (1-\gamma)\rho$$

b) The production is a failure and the outcome is bad (b)

$$\Pr(C|b) = \frac{\Pr(b|C)\phi}{\Pr(b|C)\phi + \Pr(b|I)(1-\phi)} \tag{2}$$

where

$$\Pr(b|C) = \gamma[\rho\tau_G + \alpha(1-\tau_G)] + (1-\gamma)[(1-\alpha)\tau_B + (1-\rho)(1-\tau_B)]$$

$$\Pr(b|I) = \gamma\alpha + (1-\gamma)(1-\rho)$$

Lemma 2 ***Beliefs updating after scapegoating (second round ONLY after a failure in the first round)***

After blaming (nested stage after a failure), two updates can occur *a) The "court" decides the employee was the culprit of the failure (b, e_c)*

$$\Pr(C|b, e_c) = \frac{\Pr(e_c|C, b)\phi_b^{pb}}{\Pr(e_c|C, b)\phi_b^{pb} + \Pr(e_c|I, b)(1 - \phi_b^{pb})} \tag{3}$$

where $\phi_b^{pb} = \Pr(C|b)$

$$\Pr(e_c|C, b) = \gamma[x\tau_G + z(1 - \tau_G)] + (1 - \gamma)[x\tau_B + z(1 - \tau_B)]$$

$$\Pr(e_c|I, b) = y$$

b) The "court" decides the superior was the culprit of the failure (b, s_c)

$$\Pr(C|b, s_c) = \frac{\Pr(s_c|C, b)\phi_b^{pb}}{\Pr(s_c|C, b)\phi_b^{pb} + \Pr(s_c|I, b)(1 - \phi_b^{pb})} \tag{4}$$

where

$$\Pr(s_c|C, b) = \gamma[(1-x)\tau_G + (1-z)(1-\tau_G)] + (1-\gamma)[(1-x)\tau_B + (1-z)(1-\tau_B)]$$

$$\Pr(s_c|I, b) = (1 - y)$$

Corollary 3 *In the efficient equilibrium.*
a) After production, $\Pr(C|g) > \Pr(C) > \Pr(C|b)$ *for all* $\phi \in (0, 1)$
b) After scapegoating:
If $x > y$, $\Pr(C|b, e_c) > \Pr(C|b) > \Pr(C|b, s_c)$ *for all* $\phi \in (0, 1)$
If $x = y$, $\Pr(C|b, e_c) = \Pr(C|b) = \Pr(C|b, s_c)$ *for all* $\phi \in (0, 1)$
If $x < y$, $\Pr(C|b, e_c) < \Pr(C|b) < \Pr(C|b, s_c)$ *for all* $\phi \in (0, 1)$

Proof. In the efficient equilibrium $(\tau_G(\phi) = \tau_B(\phi) = 1)$
a) From Lemma 1, $\Pr(C|g) > \Pr(C) > \Pr(C|b)$ because $\gamma(1-\rho)+(1-\gamma)\alpha > \gamma(1 - \alpha) + (1 - \gamma)\rho$ since $\alpha > \rho$ by assumption.
b) From Lemma 2, $\Pr(e_c|C, b) = x$ and $\Pr(e_c|I, b) = y$. Depending on the ordering between x and y, the relations in the Corollary hold. ■

Definition 4 *A Markov perfect equilibrium*[9] *is: Probabilities of hiring experts both in good and bad times (τ_G and τ_B), blaming intensities (x, y and z), probabilities "consumers" assign to receiving a good outcome ($p(\phi)$) given a reputation prior $\phi = \Pr(C)$, and posterior beliefs $\varphi = \Pr(C|S, \phi)$ where S are the three possible states $S \in \{g; (b, e_c); (b, s_c)\}$, such that:*

[9] We require behavior to be Markov in order to eliminate equilibriums that depend on implausible degrees of coordination between the superior behavior and "consumers" belief's about that superior behavior. (See discussion in Mailath and Samuelson, 1998).

1) Delegation decisions by competent superiors
$\tau_G(\phi)$ (in good times) and $\tau_B(\phi)$ (in bad times) maximize a value function $V(\phi)$ for all possible reputation values ϕ

2) Blaming intensities by superiors
x, y and z maximize the value function $V(\phi)$ 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)

$$p(\phi) = \Pr(g|\phi) = \Pr(g|C)\phi + Pr(g|I)(1-\phi) \tag{5}$$

4) Beliefs about competence (updated using Bayes rule).
***a)** Update after a good outcome (g)*

$$\varphi(\phi|g) = \phi_g = (1-\lambda)\Pr(C|g) + \lambda\theta \tag{6}$$

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

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

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

$$\varphi(\phi|b, s_c) = \phi_b^{s_c} = (1-\lambda)\Pr(C|b, s_c) + \lambda\theta \tag{8}$$

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

3 Efficient Equilibrium, Inefficient Scapegoating

This paper has a fundamental question. Does scapegoating really reduce the probability of achieving an efficient outcome?

With this question in mind we focus on the conditions for an efficient situation to be sustained as an equilibrium[10]. Considering the assumption $(\alpha - \rho) > (w_E - w_N) > 0$, efficiency is achieved when competents always hire experts,

[10] This model has multiple equilibria, including a very inefficient one that may arise without conditions, in which competents only hire nonexperts. Intuitively, if "consumers" think competents will hire nonexperts they will not update beliefs and competents will optimally prefer never to hire experts, who are more expensive workers and given beliefs, do not represent any additional benefit in terms of reputation. This is an equilibrium because superior's strategies uniquely determine the equilibrium updating rule that "consumers" must use if their beliefs are to be correct.

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 difference in wages $w_E - w_N$ has to be smaller than Δ. This cutoff is obtained both in 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 just as a benchmark to see how results differ when superiors can impunely blame workers.

As will be shown, whenever $\overline{x} > \overline{y}$ and blaming capabilities of competents are high enough (specifically, when a sufficient condition $\overline{z} \geq 1 - \frac{\rho}{\alpha}(1-\overline{x})$ holds),

$$\Delta_G^{NS} = \Delta_B^{NS} \geq \Delta_G^S \geq \Delta_B^S \tag{9}$$

These simple inequalities, which are in fact typically strict, summarize the main conclusions of the paper. Given a wage differential in the economy, the first inequality says that scapegoating makes the condition for an efficient equilibrium $\Delta \geq (w_E - w_N) > 0$ more difficult to hold. Furthermore it will be shown that $\Delta_G^{NS} = \Delta_B^{NS} > 0$, which means that without scapegoating it is always possible to find a positive wage differential close enough to zero that sustains efficiency, which is not necessarily the case under the presence of scapegoating.

The second inequality says that it is even more difficult to achieve efficiency with scapegoating in bad times than in good times. This will be called "Machiavellian Effect", a feature consistent with real examples and, as shown by the equality in (9), not captured when scapegoating is not considered as a choice.

3.1 Conditions for efficient equilibrium

As a first step we present the condition for the existence of an efficient equilibrium without scapegoating, which is not only easier to interpret but also helps to build on intuition.

In this case there are only two possible states (g and b) 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 (defined similarly to ϕ_g in equation (6)). The proof is in the Appendix.

Proposition 5 ***Efficient Equilibrium without Scapegoating***

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

$$\Delta^{NS} = \Delta_G^{NS} = \Delta_B^{NS} = \min_{\phi\in[\lambda\theta,1-\lambda(1-\theta)]} \left\{\delta(1-\lambda)[X + \delta(1-\lambda)V_f]\right\} > 0 \tag{10}$$

such that, for all $\Delta^{NS} \geq (w_E - w_N) > 0$*, the efficient pure strategy profile in which competents always hire experts is a Markov perfect equilibrium.*

where

$$V_f = \Pr(g|DE)Y_g + \Pr(b|DE)Y_b$$
$$X = (\alpha - \rho)[p(\phi_g) - p(\phi_b)]$$
$$Y_i = (\alpha - \rho)[V(\phi_{gi}) - V(\phi_{bi})] \qquad \textit{for } i \in \{g, b\}$$

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

Proposition 6 ***Efficient Equilibrium with Scapegoating***

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

a) If $\overline{x} \leq \overline{y}$, conditions for an efficient equilibrium are exactly the same as the case without scapegoating (Proposition 5).

b) If $\overline{x} > \overline{y}$, there exists a, non-necessarily positive, cutoff for each state of the world $s \in \{B, G\}$

$$\Delta_s^S = \min_{\phi \in [\lambda\theta, 1-\lambda(1-\theta)]} \left\{ \delta(1-\lambda)[X^s + \delta(1-\lambda)V_f^s] \right\} \tag{11}$$

such that, for all $\Delta_s^S \geq (w_E - w_N) > 0$, the efficient pure strategy profile in which competents always hire experts is a Markov perfect equilibrium.

where

$$V_f^s = \Pr(g|DE)Y_g^s + \Pr(b|DE)Y_b^s$$
$$X^B = (\alpha - \rho)p(\phi_g) + (1-\alpha)p(\phi_{b,DE}) - (1-\rho)p(\phi_{b,DN})$$
$$X^G = (\alpha - \rho)p(\phi_g) + \rho p(\phi_{b,DE}) - \alpha p(\phi_{b,DN})$$

being

$$\phi_{b,DE} = \overline{x}\phi_b^{e_c} + (1-\overline{x})\phi_b^{s_c} \tag{12}$$
$$\phi_{b,DN} = \overline{z}\phi_b^{e_c} + (1-\overline{z})\phi_b^{s_c} \tag{13}$$

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

$$Y_i^B = (\alpha-\rho)V(\phi_{gi}) + (1-\alpha)[\overline{x}V(\phi_{bi}^{e_c}) + (1-\overline{x})V(\phi_{bi}^{s_c})] - (1-\rho)[\overline{z}V(\phi_{bi}^{e_c}) + (1-\overline{z})V(\phi_{bi}^{s_c})]$$

$$Y_i^G = (\alpha-\rho)V(\phi_{gi}) + \rho[\overline{x}V(\phi_{bi}^{e_c}) + (1-\overline{x})V(\phi_{bi}^{s_c})] - \alpha[\overline{z}V(\phi_{bi}^{e_c}) + (1-\overline{z})V(\phi_{bi}^{s_c})]$$

A couple of features are worth noting before going to the main proposition of the paper. First, it's necessary to emphasize that the non-scapegoating case is just a particular example of the scapegoating case. When $\overline{x} \leq \overline{y}$ both cases are in fact exactly the same. When $\overline{x} > \overline{y}$, as $\overline{x}$, $\overline{y}$, $\overline{z} \to 0$ (maintaining the relation $\overline{z} \geq \overline{x} > \overline{y}$) always $\phi_{b,DN} \to \phi_{b,DE} \to \phi_b$ (as can be checked easily from equations (7), (8), (12), (13) and Lemma 2). Hence $X^B \to X^G \to X = (\alpha - \rho)[p(\phi_g) - p(\phi_b)]$, and $V_f^B \to V_f^G \to V_f$. This is the same as saying that cutoffs in all situations approach each other ($\Delta_B^S \to \Delta_G^S \to \Delta^{NS}$), or that Proposition 6 approaches Proposition 5, as the importance of blaming disappears.

Second, in the differing case ($\overline{x} > \overline{y}$), since $(w_E - w_N)$ is positive by assumption and there is no way to know the sign of Δ_s^S, it can only be said that whenever $\Delta_s^S < 0$, no wage differential $(w_E - w_N)$ can possibly support an efficient equilibrium. Even when in the absence of scapegoating there is always a positive wage differential that supports an efficient equilibrium, this is not necessarily true under scapegoating possibilities. This naturally goes in the proposed direction that scapegoating is harmful for efficiency, which will be formalized in the next subsections.

In the remainder of the paper, and unless stated otherwise, when referring to the scapegoating case we will be referring specifically to the case where $\overline{x} > \overline{y}$, the only interesting situation in which scapegoats constitutes a problem.

3.2 Scapegoating Inefficiency

Next, the most important conclusion of the paper, the negative impact of scapegoating on the probability to sustain efficiency, is derived. The strategy is to prove that conditions for an efficient equilibrium with scapegoating when $\overline{x} > \overline{y}$ (from Proposition 6) are more difficult to hold than without scapegoating (from Proposition 5).

Proposition 7 ***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 better blaming capabilities than inepts ($\overline{x} > \overline{y}$). It is always possible to find a $\overline{z} \geq z^* = 1 - \frac{\rho}{\alpha}(1 - \overline{x})$ such that the range of wage differentials $w_E - w_N > 0$ 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 enough since in the "Machiavellian Effect" Theorem (Proposition 10) ahead it will be shown that always $\Delta_G^S \geq \Delta_B^S$. This proof is based on the simpler case in which scapegoating is not a possibility in the future, only in the current period. The conclusion for the more general case does not varies but it is characterized by awkward statements (shown in the Appendix). We consider only the relevant case in which $\overline{x} > \overline{y}$ and there is a separating blaming equilibrium such that $\phi_b^{e_c} > \phi_b > \phi_b^{s_c}$

We will proceed in three steps. First we will show that $\phi_{b,DN} \geq \phi_{b,DE}$, second that $\phi_{b,DE} \geq \phi_b$ (as defined in Proposition 6) 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: *(* $\phi_{b,DN} \geq \phi_{b,DE}$ *)*
Consider beliefs about decision rules in the efficient equilibrium ($\tau_G(\phi) = \tau_B(\phi) = 1$), from (13), (7) and (8),

$$\phi_{b,DN} = \overline{z}\phi_b^{e_c} + (1-\overline{z})\phi_b^{s_c}$$

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

and,from (12), (7) and (8),

$$\phi_{b,DE} = \overline{x}\phi_b^{e_c} + (1-\overline{x})\phi_b^{s_c}$$

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

where $\phi_b^{pb} = \Pr(C|b)$, as defined in Lemma 1 and 2[11].
Subtracting both expressions.

$$\phi_{b,DN} - \phi_{b,DE} = \frac{(1-\lambda)\phi_b^{pb}(1-\phi_b^{pb})(\overline{z}-\overline{x})(\overline{x}-\overline{y})}{\overline{y}(1-\overline{y}) + \phi_b^{pb}(\overline{x}-\overline{y})[1-2\overline{y}-\phi_b^{pb}(\overline{x}-\overline{y})]} \tag{14}$$

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

Step 2: *(* $\phi_{b,DE} \geq \phi_b$ *)*
Subtracting $\phi_b = (1-\lambda)\Pr(C|b) + \lambda\theta$ from (12).

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

which implies

$$\phi_{b,DE} - \phi_b = \frac{(1-\lambda)\phi_b^{pb}(1-\phi_b^{pb})^2(\overline{x}-\overline{y})^2}{\overline{y}(1-\overline{y}) + \phi_b^{pb}(\overline{x}-\overline{y})[1-2\overline{y}-\phi_b^{pb}(\overline{x}-\overline{y})]} \tag{15}$$

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

Step 3: *(* $\Delta^{NS} \geq \Delta_G^S$ *for all* $\phi \in (0,1)$ *)*
Given equations (10) and (11), it's sufficient to show the following two claims.

Claim 1) $X \geq X^G$ *for all* ϕ*.*
Subtracting these expressions

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

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

which is non-negative since $\alpha > \rho$ by assumption and $p(\phi_{b,DN}) \geq p(\phi_{b,DE})$ for all feasible ϕ by step 1 (equation 14) and monotonicity of $p(\phi)$.

Claim 2) $V_f \geq V_f^G$ for all ϕ.
Subtracting these expressions

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

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

$$Y_i - Y_i^G = (\alpha\overline{z} - \rho\overline{x})[V(\phi_{bi}^{e_c}) - V(\phi_{bi}^{s_c})] - (\alpha - \rho)[V(\phi_{bi}) - V(\phi_{bi}^{s_c})]$$

is non-negative.

Because of the monotonicity of $V(\phi)$ in ϕ and since, by equation (7) and Corollary 3, $\phi_b^{e_c} \geq \phi_b$, then $V(\phi_{bi}^{e_c}) \geq V(\phi_{bi})$. A sufficient condition for non-negativity is then $(\alpha\overline{z} - \rho\overline{x}) \geq (\alpha - \rho)$. This condition is fulfilled whenever,

$$\overline{z} \geq z^* = 1 - \frac{\rho}{\alpha}(1 - \overline{x}) \tag{16}$$

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 $\overline{z} \geq z^*$ holds, regardless of the value function's behavior, the likelihood of having an efficient situation is the highest without scapegoating or, which is the same in this model, when the credibility assigned to evidence presented by superiors against subordinates is low in general. ■

The intuition behind this general result is that reports after a failure when $\overline{x} > \overline{y}$ represent a way for competents to further signal their competence. If this is the case, competents can exploit differences in the blaming capacity 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 $(\overline{x} - \overline{y})$, which reduces reputation losses after bad results. Second, by the assumed difference between blaming experts and nonexperts $(\overline{z} - \overline{x})$, which may introduce an additional gain from hiring nonexperts (additional to the lower cost that w_N represents), by increasing the probability that the "court" assigns the fault of the failure to the employee.

Now, it is important to put into context the sufficient condition for the inefficiency of scapegoating, $z \geq z^*$. This is relevant because we do not know the behavior of the value function. But, for example, if the value function were linear, both Y_g and Y_b would behave exactly as X and scapegoating would always imply inefficiency, regardless of the specific value of $\overline{z}$.

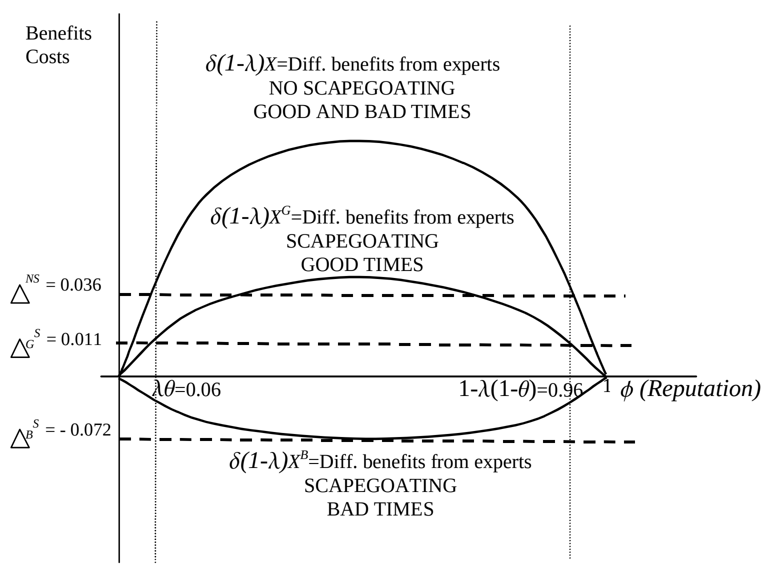
As can be seen z^* depends on $\frac{\rho}{\alpha}$, a measure of the relative capability of experts to achieve good outcomes in production when compared to nonexperts. For example, if hiring experts almost guarantees success in good times $(\rho \to 0)$, then $z^* \to 1$. At the other side, if hiring experts does not add a lot to the probability of success $(\rho \to \alpha)$, then $z^* \to \overline{x}$.

This implies that the sufficient condition $\overline{z} \geq z^*$ is more difficult to hold when hiring experts is really beneficial from a productive point of view, which means superiors can signal their competence directly in the first stage, without the need to go to "court". On the other hand, when hiring experts does not make an important difference in production, competents tend to rely more on the use of scapegoating to signal competence, leading heavily towards inefficiency.

Conditions for an efficient equilibrium in the three cases discussed previously, without scapegoating (both in good and bad times, given by Δ^{NS}), with scapegoating in good times (given by Δ_G^S) and with scapegoating in bad times (given by Δ_B^S) can be easily seen in Figure 1. In this case we assumed that $V_f = V_f^G = V_f^B = 0$, which easy computations conservatively biasing results in favor of hiring experts. Even in this conservative situation, not allowing for scapegoating (*not asking why things went wrong!*) increases the incentives to efficiently hire experts.[12].

Figure 1

Example of conditions for Efficient Equilibrium



[12]Parameters used: $\lambda = 0.1$, $\theta = 0.6$, $\delta = 0.99$, $\rho = 0.1$, $\alpha = 0.4$, $\gamma = 0.5$, $\overline{y} = 0.15$, $\overline{x} = 0.3$ and $\overline{z} = 0.85$. The sufficient condition from equation (16) holds because in this case $\overline{z} > z^* = 0.825$

3.3 The intuition lying behind scapegoating inefficiency

The whole action in previous theorems and proofs comes from the comparison of reputation competents expect to obtain from hiring experts as opposed to hiring nonexperts.

Without scapegoating, the reputation conditional on the first round's results (i.e. conditional on good and bad outcomes) is known and given by ϕ_g after a success (as defined in equation (6)) and ϕ_b after a failure (defined also in a similar way than (6) but using equation (2) instead).

With scapegoating, while the expected reputation after a good outcome is also independent of the hiring decision, ϕ_g (because there is no second stage after a success), the expected reputation after bad results depends on the hiring decision. It matters for blaming whether the superior works with an expert or with a nonexpert since experts are harder to be blamed successfully.

For example, the expected reputation after a failure when hiring experts is given by,

$$E(\Pr(C|b)|DE) = \phi_{b,DE} = \Pr(e_c|DE)\varphi(\phi|b, e_c) + \Pr(s_c|DE)\varphi(\phi|b, s_c)$$

which is the expression in equation (12). Similarly, the expected reputation after a failure when hiring nonexperts, $\phi_{b,DN}$, was defined in equation (13).

The differential gains in reputation expected from good results in production can be seen as a measure of the incentives to hire experts, since this decision increases the probabilities of success. These gains can be represented by $(\phi_g - \phi_b)$ without scapegoating (regardless of the hiring decision), by $(\phi_g - \phi_{b,DE})$ with scapegoating if the decision is to hire experts and by $(\phi_g - \phi_{b,DN})$ with scapegoating if the decision is to hire nonexperts.

Hence, to understand how incentives to efficiently hire experts behave we need to understand how $\phi_{b,DN}$, $\phi_{b,DE}$ and ϕ_b relate to each other.

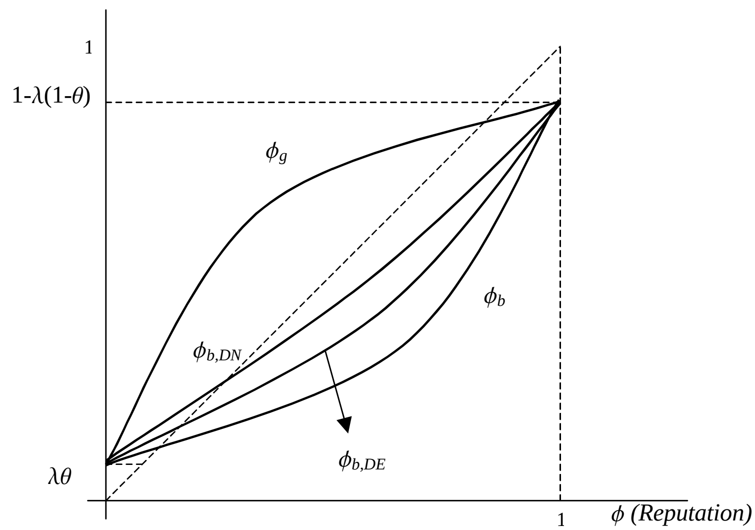
As was shown in steps 1 and 2 of Proposition 7's proof, there is a clear ordering between these expressions when $\overline{z} > \overline{x} > \overline{y}$.[13].

$$\phi_g > \phi_{b,DN} > \phi_{b,DE} > \phi_b \tag{17}$$

Graphically, Figure 2 shows the updated reputation value for each case and each possible reputation prior ϕ.

[13]For this comparison, do not consider extreme cases for $\overline{z}$, $\overline{x}$ and $\overline{y}$. This is because ϕ_g may be even smaller than $\phi_{b,DE}$. These cases will be discussed later and do not change, just reinforce, the results.

Figure 2
Expected reputation after the first round
With and without scapegoating



It is easy to check from the equation for each case (Lemmas 1 and 2 and equations (14) and (15)) that all possible beliefs' updates 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 shown by relation (17) and Figure 2.

Going back to equations (14) and (15), while the difference $(\phi_{b,DE} - \phi_b)$ can be interpreted as the reduction in the incentives to hire experts, the difference $(\phi_{b,DN} - \phi_{b,DE})$ can be seen as the increase in the incentives to hire nonexperts.

While the expression $(\overline{x} - \overline{y})$ in the numerator of equations (14) and (15) shows the magnitude of reputation maintenance due to scapegoating, the expression $(\overline{z} - \overline{x})$ on the numerator of equation (14) shows the additional benefits from hiring nonexperts by taking advantage of the chances to blame them after a failure. When compared to the case in which blaming is not an option, scapegoating reduces the incentives to hire experts by the gap $(\phi_{b,DE} - \phi_b)$ and increases the incentives to hire nonexperts by the gap $(\phi_{b,DN} - \phi_{b,DE})$.

Not only is the ordering clear for all $\phi \in (0, 1)$ but also the impact of blaming capacities' gaps $(\overline{x} - \overline{y})$ and $(\overline{z} - \overline{x})$ on equations (14) and (15). The following lemmas show that both $(\phi_{b,DN} - \phi_{b,DE})$ and $(\phi_{b,DE} - \phi_b)$, not only are positive (as shown in steps 1 and 2 in Proposition 7's proof) but also depend positively on the blaming abilities' gaps.

Lemma 8 *The difference between expected reputation after a failure from hiring experts versus hiring nonexperts ($\phi_{b,DN} - \phi_{b,DE}$) is non-decreasing in ($\overline{z} - \overline{x}$) nor in ($\overline{x} - \overline{y}$)*

Proof. a) Taking the derivative of expression ($\phi_{b,DN} - \phi_{b,DE}$) in (14) with respect to the difference ($\overline{z} - \overline{x}$)

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

b) Taking the derivative of expression ($\phi_{b,DN} - \phi_{b,DE}$) in (14) with respect to the difference ($\overline{x} - \overline{y}$)

$$\frac{\partial[\phi_{b,DN} - \phi_{b,DE}]}{\partial(\overline{x} - \overline{y})} = \frac{(1-\lambda)\phi_b^{pb}(1-\phi_b^{pb})(\overline{z} - \overline{x})[\overline{y}(1-\overline{y}) + \phi_b^{pb}(\overline{x} - \overline{y})^2]}{\left[\overline{y}(1-\overline{y}) + \phi_b^{pb}(\overline{x} - \overline{y})[1 - 2\overline{y} - \phi_b^{pb}(\overline{x} - \overline{y})]^2\right]^2} \geq 0$$

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

Lemma 9 *The difference between expected reputation after a failure in cases with and without scapegoating, ($\phi_{b,DE} - \phi_b$) is non-decreasing in ($\overline{x} - \overline{y}$)*

Proof. For this proof just consider the difference $\phi_{b,DE} - \phi_b$ in (15) since, as shown in Lemma 8, $\frac{\partial[\phi_{b,DN} - \phi_{b,DE}]}{\partial(\overline{x} - \overline{y})} \geq 0$. Taking derivatives of ($\phi_{b,DE} - \phi_b$) with respect to ($\overline{x} - \overline{y}$)

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

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

The difference in the blaming abilities between competents and inepts ($\overline{x}-\overline{y}$) basically measures the drop in expected reputation that, thanks to scapegoating, does NOT occur after a failure. Hence, an increase in ($\overline{x} - \overline{y}$) not only reduces the incentives to hire experts (by increasing $\phi_{b,DE} - \phi_b$) but also makes more beneficial to hire nonexperts (by increasing $\phi_{b,DN} - \phi_{b,DE}$ as well).

Similarly, the difference in the abilities between blaming experts and nonexperts ($\overline{z} - \overline{x}$) measures the greater probability of having a positive "court" decision against employees if hiring nonexperts. Hence an increase in ($\overline{z} - \overline{x}$) makes even more beneficial to hire nonexperts (by further increasing $\phi_{b,DN} - \phi_{b,DE}$).

3.4 Machiavellian Effect

The next proposition shows that in bad times an efficient outcome is more difficult to obtain than in good times. Reputation concerns and scapegoating rationalize this "Machiavellian Effect" since in bad times superiors are more worried about potential reputation losses rather than potential reputation gains.

Proposition 10 ***"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 competents have better blaming capabilities than inepts ($\overline{x} > \overline{y}$). If there exist some $w_E - w_N > 0$ such that competents decide to hire experts in bad times, then they also decide to hire experts in good times, while the contrary is not necessarily true.

Proof. In order to prove how blaming makes conditions for efficient equilibrium more difficult to fulfill in bad times than in good times 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 (11) it suffices to show the following two claims,

Claim 1) $X^G \geq X^B$ for all ϕ.
Subtracting these expressions

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

which is non-negative since $\alpha+\rho < 1$ by assumption; $p(\phi)$ is monotonic in ϕ and by equation (14) $\phi_{b,DN} \geq \phi_{b,DE}$ in the separating blaming equilibrium where $\overline{x} > \overline{y}$ and $\overline{z} \geq \overline{x}$.

Claim 2) $V_f^G \geq V_f^B$ for all ϕ
Subtracting these expressions

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

which is non-negative because $Y_i^G - Y_i^B = (1-\alpha-\rho)(\overline{z}-\overline{x})[V(\phi_{bi}^{e_c}) - V(\phi_{bi}^{s_c})] \geq 0$ for $i \in \{g,b\}$. This is because $\alpha+\rho < 1$ and $\overline{z} \geq \overline{x}$ by assumption, $V(\phi)$ is monotonic in ϕ and $\phi_b^{e_c} \geq \phi_b^{s_c}$ by Corollary 3 and separating equilibrium ($\overline{x} > \overline{y}$).

Assuming scapegoating both now and in the future the conclusion and proof are exactly the same.

Hence, in good times experts are hired for a wider range of wage differentials ($w_E - w_N$) than in bad times. This does not imply a positive cutoff Δ_s^S, but it does imply it is more likely to have $\Delta_G^S > 0$ rather than $\Delta_B^S > 0$ and then, that efficiency be achieved in good times but not in bad times. This is what we called "Machiavellian Effect". ■

The intuition behind the "Machiavellian Effect" is that, even when in good and bad times differences in probabilities to obtain a good outcome ($\alpha - \rho$) are the same[14], the probability of having a bad outcome is greater in bad times than in good times (in our model, greater than half). Under the presence of scapegoating this is important because of the possibility to avoid a big reduction in reputation if a failure in fact occurs, hence making more attractive the hiring of nonexperts who can be blamed easily.

3.5 Discussion of results

The relevance of the blaming report for efficiency resides both in its value to competents to further signal their competence and in its unproductiveness to "consumers". Since by assumption superiors only care about reputation, decisions will react more to activities that allow a better competence signaling. If those activities only occur after particular situations, such as scapegoating only happens after failures, superiors will make decisions trying to achieve those circumstances, even when detrimental to activities that really matter to society.

To put the results into context, as is generally the case, it is useful to refer to extreme cases. Assume for a second $\overline{y} = 0$ and $\overline{x} = \overline{z} = 1$. This means inepts cannot convince anybody about the blame of employees while competents can always blame others convincingly in front of a "court". In this situation, if "consumers" see that after a failure the "court" decides against subordinates, they learn for sure the superior is competent, increasing immediately the reputation. Hence blaming is better than production for competents to signal their competence. In fact they will prefer to have a failure in order to show their capabilities more effectively through "court" rather than through production performance, a possibility clearly not allowed after a success.

In this very extreme example, competents will never hire experts because nonexperts are not only less expensive but also increase the probability of going to "court". That's why the position of a nested activity in the reputation game is very important to understand its outcome in terms of efficiency.

Naturally, the previous extreme example is consistent with the case $\overline{x} > \overline{y}$. But what happens if $\overline{x} \leq \overline{y}$? As shown formally, in this case blaming is useless to signal competence and competents will not behave differently than without scapegoating. To see it clearly, consider now the other extreme in which inepts can always blame effectively ($\overline{y} = 1$) while competents have zero blaming capacity ($\overline{x} = 0$). In this case, as in all other cases in which $\overline{x} \leq \overline{y}$, it is not credible that inepts will use all their abilities to convince courts about the employee's fault. Why would inepts like to blame workers so effectively if by doing so they just signal their ineptitude? To adjust his blaming efforts downwards (say not presenting proofs, burning evidence against employees, etc.) and to be confused with competents seems to be a better strategy. Blaming is an activity, as are many others, where the success probability can be easily adjusted downwards (just being lazy) but not upwards.

[14] This is just an assumption to clarify the "Machiavellian Effect" as much as possible. To assume otherwise does not change the main conclusion.

Since inepts always want to be confused with competents, in cases where they are not worse at blaming than competents ($\overline{x} \leq \overline{y}$), they can decrease y until it reaches $\overline{x}$ by reducing blaming intensities and efforts to collect evidence. Potentially, competents can always be imitated by inepts, not being an equilibrium the use of the blaming report to impose a new updating round after production.

Therefore, only when competents outperform in the blaming activity ($\overline{x} > \overline{y}$), inepts cannot do anything to achieve competent levels $\overline{x}$ (In fact, that's exactly why they are inepts!). Naturally they would like to be confused with competents but they simply cannot be so good. Contrarily, competents want to separate as much as possible from inepts ($\overline{x}$), who will just do the best they can ($\overline{y}$). This is the situation where scapegoating introduces a real difference in the model, leading to inefficiencies.

4 Ways to increase the likelihood of an efficient equilibrium (without spending more money!).

Up to this point we conclude that scapegoating generates inefficiencies because of its location after bad results in the nested reputation game. Then, a natural question arises. What happens with activities nested after successes?

Examples of these kinds of situations can be found in abundance in real life and, as will be shown, increase the likelihood to reach efficiency when competents are not only better at producing but also at those nested activities.

4.1 Examples of nested activities after successes

Many activities that fit into the characterization of unproductive nested stages after successes can be found in different spheres.

A lot of examples belong to the sporting arena. Typical cases are All-Star Games, exhibition matches performed by the best players in the respective sport league. These players are often chosen by a popular vote of fans, and the game, which mostly occurs at the halfway point of the regular season, is basically unproductive given there is no gain or loss whether or not the competitors are victorious. All-Star Games are organized by a huge variety of sport leagues such as, MLB, NHL, NBA, MLS, NLL, NFL, NASCAR, etc.

Players who participate are those very successful in the respective period and games, even when unproductive from the league and fans' standpoint, represent additional ways in which the best players can further signal their competence in front of a greater mass of people than the one attending regular season matches.

Another example is the formation of national teams and the organization of international championships (such as the Soccer World Cup) in almost all existing sports. Only the best players are chosen to be part of national teams and, even when the objective is productive, it represents an additional way successful players can signal their competence in front of the whole world. Typically players are not paid to participate in these events but, after their conclusion, the best ones generally obtain better contracts and are hired by important teams.

Examples of nested activities after successes can be found in some organizations and corporations as well. In many hierarchies, after a division outperforms relative to the rest of the organization, additional funds and decision responsibilities are assigned to it. In this way the superiors of those agencies have alternative ways to signal their competence, even when the new responsibilities are not necessarily relevant to the normal functioning of the division or to the goals of the organization.

NFI Research, a global research firm based in New Hampshire, conducted a survey that shows that more than half the managers prefer to be recognized for a good job by an "increase in responsibility and opportunities to attend external events". Chuck Martin, NFI's chairman, wrote in Darwin Magazine: "Businesspeople are not just looking for the employee-of-the-month parking spot. In fact, the last things they want are trophies, awards...or even time off". In the same vein, Kraft Bell in a management report[15] stated, "People want to be empowered to have more involvement and collaboration in decisions".

Additional cases can also be discovered in academic environments. Top researchers and professors are typically invited to participate in round tables and plenary sessions at professional meetings. These activities, only opened to professionals very successful on their fields, can be considered as additional ways available to distinguished academicians to further signal their competence in front of their counterparts. In fact this may also be part of the reason congresses and conferences with contributed sessions exist.

In all these academic cases research constitutes the productive and important activity. The following stage (the conference, for example), even when productive in itself, is also very useful to further signal the competence of the best.

Another environment where it is possible to find activities nested after successes is show business. Successful actors, actresses, music and movie stars are invited to participate in TV interviews. Even when not paid, being on famous TV shows represents additional ways to signal competence by exhibiting charisma or being funny.

A final interesting case is the new performance-based scheme that New York is using to improve schools' quality in which "head teachers of successful schools receive cash bonuses worth up to $15,000 (to develop new activities)". Furthermore "The New York scheme...is intended to promote successful staff and also to make heads more directly accountable for their performance"[16] This NY policy is a combination of more funds and activities principals can use to signal their competence after successes and a reduction in scapegoating possibilities after failures. This integral scheme goes exactly in the direction proposed by this paper.

[15] Gartner, G2, August 2003 Report.

[16] BBC News, Wednesday, 30 May, 2001

4.2 Efficiency of nested activities after successes.

The model can be easily reinterpreted and modified to introduce nested activities after successes. Assume that instead of an unproductive activity after a failure, as was the case with scapegoating, the game is characterized by an unproductive activity nested after successes (think about any of the previous examples). The structure of probabilities, timing and parameters for the production stage have the same interpretation as before. Hence, without nested activities, competents compare the expected reputation gains from improving the chances of getting a good result in production with the greater costs from hiring experts.

The difference appears in the second stage. After bad results the game is over but after good results there is a nested activity, which at the time can be a success (g, s) or a failure (g, f) (these have basically the same spirit than (b, e_c) and (b, s_c) in the original model). Using the same notation as before we can write $\Pr(s|C, g, DE) = x_g \in [0, \overline{x_g}]$ and $\Pr(f|C, g, DN) = y_g \in [0, \overline{y_g}]$, which means some differences in the capabilities of being successful at the nested activity may exist[17].

For example, if $\overline{x_g} > \overline{y_g}$, hiring experts not only increases the probability of being successful at production but also at the additional nested activity. Superiors may choose the probabilities of being successful at the nested activity by eventually boycotting the worker efforts and actions. If there is no boycott, the probability of success will be the maximum (say $x_g = \overline{x_g}$ if the employee is an expert) and a maximum boycott intensity by the superior eliminates the probability of success in the nested activity ($x_g = 0$). This decision about the boycotting intensity has the same logic that the decision about the blaming intensity in the original model with scapegoating.

The equilibrium definition in this environment is the same as before with the difference that the three possible updating (in place of equations (6)-(8)) are now ϕ_b (after a bad outcome), ϕ_g^s (after successes both in the first and second rounds) and ϕ_g^f (after success in production and failure in the nested unproductive activity).

Therefore, we need to restate Lemma 2 in the following way: After a good outcome, two updates can occur.

a) If the nested activity is successful (g, s)

$$\Pr(C|g, s) = \frac{\Pr(s|C, g)\phi_g^{pb}}{\Pr(s|C, g)\phi_g^{pb} + \Pr(s|I, g)(1 - \phi_g^{pb})} \tag{18}$$

where

$$\Pr(s|C, g) = \gamma[x_g\tau_G + y_g(1 - \tau_G)] + (1 - \gamma)[x_g\tau_B + y_g(1 - \tau_B)]$$

[17] Many of the examples discussed in 4.1 do not need delegation because there is no scapegoating. 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, even the same in spirit, is different in that superiors would need to choose the effort level both before the first and second rounds and pay twice the effort costs. Introducing this modification does not substantially change the main conclusion, though.

$$\Pr(s|I,g) = y_g$$

b) If the nested activity is unsuccessful (g,f)

$$\Pr(C|g,f) = \frac{\Pr(f|C,g)\phi_g^{pb}}{\Pr(f|C,g)\phi_g^{pb} + \Pr(f|I,g)(1-\phi_g^{pb})} \tag{19}$$

where

$$\Pr(f|C,g) = \gamma[(1-x_g)\tau_G + (1-y_g)(1-\tau_G)] + (1-\gamma)[(1-x_g)\tau_B + (1-y_g)(1-\tau_B)]$$

$$\Pr(f|I,g) = (1-y_g)$$

Similarly to Corollary 3 it's 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 $\Pr(C|g,s) > (<) \Pr(C|g) > (<) \Pr(C|g,f)$ if $x_g > (<) y_g$

As in the scapegoating situation, only when $\overline{x_g} > \overline{y_g}$ (i.e. when hiring experts increases the probability of success both in production and in nested activities) may the nested stage generate a new reputation updating and affect efficiency[18]. Therefore, in what follows, unless stated otherwise, we consider only the relevant case in which $\overline{x_g} > \overline{y_g}$.

As in equations (14) and (15) it's also possible in this context to define expected reputation after good results in case of hiring experts and in case of hiring nonexperts. Equation (14) could be restated in this environment as,

$$\phi_{g,DE} - \phi_{g,DN} = \frac{(1-\lambda)\phi_g^{pb}(1-\phi_g^{pb})(\overline{x_g} - \overline{y_g})^2}{\overline{y_g}(1-\overline{y_g}) + \phi_g^{pb}(\overline{x_g} - \overline{y_g})[1 - 2\overline{y_g} - \phi_g^{pb}(\overline{x_g} - \overline{y_g})]} > 0 \tag{20}$$

and equation (15) could be restated as,

$$\phi_{g,DE} - \phi_g = \frac{(1-\lambda)\phi_g^{pb}(1-\phi_g^{pb})^2(\overline{x_g} - \overline{y_g})^2}{\overline{y_g}(1-\overline{y_g}) + \phi_g^{pb}(\overline{x_g} - \overline{y_g})[1 - 2\overline{y_g} - \phi_g^{pb}(\overline{x_g} - \overline{y_g})]} > 0 \tag{21}$$

In this case it's also informative to obtain the difference $\phi_{g,DN} - \phi_g$

$$\phi_{g,DN} - \phi_g = -\frac{(1-\lambda)(\phi_g^{pb})^2(1-\phi_g^{pb})(\overline{x_g} - \overline{y_g})^2}{\overline{y_g}(1-\overline{y_g}) + \phi_g^{pb}(\overline{x_g} - \overline{y_g})[1 - 2\overline{y_g} - \phi_g^{pb}(\overline{x_g} - \overline{y_g})]} < 0 \tag{22}$$

A clear ordering exists among these expressions when $\overline{x_g} > \overline{y_g}$ (and a separating equilibrium in the second stage arises),

$$\phi_{g,DE} > \phi_g > \phi_{g,DN} > \phi_b \tag{23}$$

[18]If $\overline{x_g} \leq \overline{y_g}$ inepts prefer to boycott 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 3.5).

such that Figure 3 (similar to Figure 2) can be drawn.

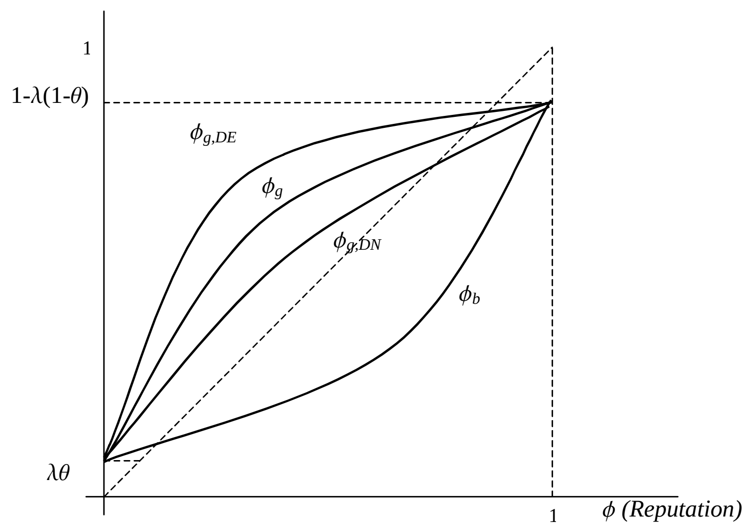
Hence hiring experts increases expected reputation after good results ((21) is positive) while hiring nonexperts in fact decreases it ((22) is negative). The final effect is an increase in gains from hiring experts.

The obvious difference between nested activities after failures (such as scapegoating) and the situation explained here is that, while the former reduces the expected reputation gains from successes (from $(\phi_g - \phi_b)$ to $(\phi_g - \phi_{b,DE})$) and decreases the incentives to hire experts, the later increases the expected reputation gains from successes (from $(\phi_g - \phi_b)$ to $(\phi_{g,DE} - \phi_b)$) and increases the incentives to efficiently hire experts. Even more, while the former increases the incentives to hire nonexperts (by $(\phi_{b,DN} - \phi_{b,DE})$), the later decreases them (by $(\phi_{g,DE} - \phi_{g,DN})$).

Figure 3

Expected reputation after the first round

With and without nested stages after successes



While Figure 3 delivers the basic intuition that sustains the efficiency of nested activities after good results when $\overline{x_g} > \overline{y_g}$, formal proofs are sketched in the Appendix (following similar strategies than Propositions 6 and 7's proofs for the scapegoating case). In this way it is possible to say that whenever the sufficient condition

$$\overline{x_g} > x_g^* = 1 - \frac{\rho}{\alpha}(1 - \overline{y_g}) \tag{24}$$

holds, the unproductive stage after a good outcome will always increase the likelihood of achieving an efficient equilibrium.

This condition means that, when experts tend to assure a production success ($\rho \to 0$), the production stage is a very powerful way to signal competence and the nested stage is not important to achieving efficiency. Contrarily, if experts do not add anything to the probability of being successful at producing ($\rho \to \alpha$), the first stage does not allow competents to signal their competence at all, being the sufficient condition just $\overline{x_g} > \overline{y_g}$ and the nested activity a powerful means of increasing efficiency. Recall that this is the same sufficient condition expressed in (16), with $\overline{z}$ and $\overline{x}$ replaced by $\overline{x_g}$ and $\overline{y_g}$ respectively.

The key here is not only to introduce nested stages after good results but also that those activities be sufficiently better performed by competents when working with experts than by inepts who can only work with nonexperts.

The nested reputation idea is based on the possibility that first stages, whose outcome really matters to "consumers", may not be so good for signaling competence in order to justify competents' efficient decisions. But as long as they give access to activities that create reputation more certainly, decisions that seem strange at first glance may actually make a lot of sense.

Finally, it is very interesting to note the "Machiavellian Effect" persists in this case as well, meaning there are more incentives to hire experts in good times than in bad times. The reason is similar to the one proposed in the scapegoating case, but in this situation the probability of being successful and reaching the nested activity is, in absolute terms, greater in good times, increasing even more the incentives to hire experts. The proof is also sketched in the Appendix.

In this section we have shown the importance of nested activities for influencing incentives, to understand superiors' decisions and to explain efficiency consequences. In general, the introduction of nested activities after successes is better than after failures in order to promote efficient decisions. The only condition is a positive correlation between abilities to be successful, both in the productive and not so productive nested activities.

Many real life situations can be rationalized from this point of view, which means it is not only a theoretical curiosity. Furthermore it is possible to think about very unproductive activities that may be useful to promote efficient results just by exploiting reputation concerns, without requiring costly incentives or monetary resources.[19]

5 Conclusion

Scapegoating is a common behavior in public institutions, firms, sports and even in the Army. The main problem is not the redistribution effects unfair blaming may generate, but the inefficiencies in the performance of organizations

[19] An extreme but funny example is the following one. Suppose that hiring experts is efficient but incentives from production are not enough for superiors to do it. Assume also 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 !

it may introduce. In fact, this popular idea has been the main argument for last decade's reforms designed to assign more responsibility to superiors, reducing their chances to blame subordinates.

Inefficiencies may arise because superiors are concerned about their reputation while "consumers" only care about production. Because of the imperfect information that begets reputation problems, interests are not aligned and superiors may not make the decisions preferred by society to achieve the best possible outcome at the lowest possible cost.

Scapegoating makes it even harder to achieve efficiency because it attenuates potential losses of reputation after failures, 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 easily if something goes wrong.

To formalize this idea, we 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 society only cares about the results in the first stage, it can be a good idea 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 rights of passing to the next one and signal their competence even better.

Furthermore, the model also delivers an interesting feature observed in reality, named here as "Machiavellian Effect", in which competents tend to hire experts more in good times than in bad times.

All in all, the model we developed allows one to understand how scapegoating may lead to inefficiencies (by reducing the possible losses in reputation and decreasing the incentives to hire experts), what the conditions are for this to happen (competents should be better blaming subordinates in general than inepts) and what the possible policies to deal with it are (not only reduce blaming possibilities by superiors, as proposed by new public sector reforms' movements, but also introducing nested activities after good results, which offer successful superiors additional opportunities to signal their competence).

Obviously these conclusions should be taken carefully. This model just focuses on one particular delegation motive, which is reputation, leaving out other important reasons such as specialization and scale. This is why conclusions are biased towards 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.

References

[1] Alesina, A. and G. Tabellini [2004]: "Bureacrats and politicians?" *NBER, Working Paper 10241.*

[2] Alesina, A. and G. Tabellini [2005]: "Why do politicians delegate?" *NBER, Working Paper 11531.*

[3] Bell, N. and P. Tetlock [1989]: "The intuitive politician and the assignment of blame in organizations". *Impression Management in Organizations*, Hillsdale, NJ: Erlbaum, 105-123.

[4] Bendor, A,; Glazer, A. and T. Hammond [2001]: "Theories of delegation". *Annu. Review of Political Science*, **4**, 235-69

[5] Cole, H. and P. Kehoe [1996]: "Reputation spillover across relationships: Reviving reputation models of debt" *NBER, Working Paper 5486.*

[6] Cripps, M; Mailath, G. and L. Samuelson [2004]: "Imperfect monitoring and impermanent reputations". *Econometrica*, **72**, 407-432

[7] Cripps, M; Mailath, G. and L. Samuelson [2004]: "Disappearing private reputations in long-run relationships". *mimeo, University of Pennsylvania*

[8] Dezso, C. [2004]: "Scapegoating and Firm Reputation". *mimeo, NYU.*

[9] Douglas, T. [1995]: "Scapegoats: Transferring blame". *mimeo, Routledge.*

[10] Fudenberg, D. and D. Levine [1989]: "Reputation and equilibrium selection in games with a single patient player" *Econometrica*, **57**, 759-778.

[11] Fudenberg, D. and D. Levine [1992]: "Maintaining a reputation when strategies are imperfectly observed" *Review of Economic Studies*, **59**, 561-579.

[12] Huson, M.; Malatesta, P. and R. Parrino [2004]: "Managerial succession and firm performance". *Journal of Financial Economics*, Forthcoming.

[13] Kreps, D. and R. Wilson [1982]. "Reputation and Imperfect information". *Journal of Economic Theory* **27**, 253-279

[14] Martin, J. [1997]: "Changing accountability relations: Politics, consumers and the market" *Public Management Service*, OECD**.**

[15] Mailath, G. and L. Samuelson [1998]: "Your reputation is who you're not, not who you would like to be" *CARESS, WP 98-11, U. of Pennsylvania.*

[16] Mailath, G. and L. Samuelson [2001]: "Who wants a good reputation?" *Review of Economic Studies*, **68**, 415-441.

[17] Milgrom, P. and J. Roberts [1982]: "Predation, reputation and entry deterrence". *Journal of Economic Theory* **27**, 280-312

[18] Polidano, C. [1999]: "The bureaucrat who fell under a bus: Ministerial responsibility, executive agencies and the Derek Lewis Affair in Britain". *Governance: An International Journal of Policy and Administration* **12**, 201-229.

[19] Segendorff, B. [2000]: "A Signalling theory of scapegoats". *mimeo, Stockholm School of Economics.*

[20] Winter, E. [2001]: "Scapegoats and optimal allocation of responsibility". *mimeo, The Hebrew University of Jerusalem.*

Appendix

Proof Lemma 1. Define in general a Bayesian updating rule for both cases

$$\Pr(C|i) = \frac{\Pr(i|C)\Pr(C)}{\Pr(i|C)\Pr(C) + \Pr(i|I)\Pr(I)}, \qquad \text{for } i \in \{g, b\} \tag{25}$$

being $\Pr(C) = \phi$ and $\Pr(I) = 1 - \phi$

Equation (25) can be decomposed to be expressed only in terms of parameters and decision rules

For example, for the case $i = g$, using the law of total probability

$$\Pr(g|C) = \Pr(G)\Pr(g|C, G) + \Pr(B)\Pr(g|C, B)$$

at the same time, in order to express everything in terms of exogenous probabilities, for example $\Pr(g|C, G)$ in the first term of the previous expression can also be decomposed in two parts

$$\Pr(g|C, G) = \Pr(g|C, G, DE)\Pr(DE|G) + \Pr(g|C, G, DN)\Pr(DN|G)$$

just replacing by defined probabilities (for example $\Pr(g|C, G, DE) = 1 - \rho$) and by decision rules (for example $\Pr(DE|G) = \tau_G$) in all terms, the expressions in Lemma 1 are obtained. ■

Proof Lemma 2. Define in general a Bayesian updating rule for both cases

$$\Pr(C|b, i_c) = \frac{\Pr(i_c|C, b)\Pr(C|b)}{\Pr(i_c|C, b)\Pr(C|b) + \Pr(i_c|I, b)\Pr(I|b)}, \qquad \text{for } i \in \{e, s\} \tag{26}$$

being $\Pr(C|b) = \phi_b^{pb}$ as defined in Lemma 1 (the superscript pb holds for "pre-blaming"). By definition $\Pr(I|b) = 1 - \phi_b^{pb}$

Equation (26) can be decomposed to be expressed only in terms of parameters and decision rules

For example, for the case $i = e_c$, using the law of total probability

$$\Pr(e_c|C, b) = \Pr(G)\Pr(e_c|C, b, G) + \Pr(B)\Pr(e_c|C, b, B)$$

at the same time, in order to express everything in terms of exogenous probabilities, for example $\Pr(e_c|C, b, G)$ in the first term of the previous expression can also be decomposed in two parts

$$\Pr(e_c|C, b, G) = \Pr(e_c|C, b, G, DE)\Pr(DE|G) + \Pr(e_c|C, b, G, DN)\Pr(DN|G)$$

Replacing by decision rules (for example $\Pr(DE|G) = \tau_G$) and considering, as an example and just to save notation, the equilibrium case in which always "consumers" correctly predict blaming intensities (for example $\Pr(e_c|C, b, G, DE) = x$), expressions in Lemma 2 are obtained. ■

Proof Proposition 5. Fix ϕ and suppose an efficient situation (i.e., the competent superior always chooses to hire an expert both in good and bad times ($\tau_G(\phi) = \tau_B(\phi) = 1$)). Under results from Corollary 3, given any state $s \in \{B, G\}$, 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 $s = B$, competent's value function when hiring experts is,

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

$$V(\phi, DE) = p(\phi) - w_E + \delta(1-\lambda)[\alpha V(\phi_g) + (1-\alpha)V(\phi_b)]$$

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

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

Thus

$$V(\phi, DE) - V(\phi; DN) = -(w_E - w_N) + \delta(1-\lambda)[X]$$
$$+\delta^2(1-\lambda)^2\{\Pr(g|DE)Y_g + \Pr(b|DE)Y_b\}$$

where

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

and $\Pr(g|DE) = \Pr(g|DE, G)\Pr(G) + \Pr(g|DE, B)\Pr(B)$

Then

$$\Pr(g|DE) = \alpha + \gamma(1 - \rho - \alpha)$$

and, in the same vein,

$$\Pr(b|DE) = (1-\alpha) - \gamma(1 - \rho - \alpha)$$

Working with the condition $V(\phi, DE) - V(\phi; DN) \geq 0$, it's necessary that cost differences $(w_E - w_N)$ fulfill

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

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

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 $\phi \in [\lambda\theta, 1 - \lambda(1-\theta)]$

$$\Delta_B^{NS} = \min_{\phi \in [\lambda\theta, 1-\lambda(1-\theta)]} \left\{\delta(1-\lambda)[X + \delta(1-\lambda)V_f]\right\} \qquad (27)$$

If $s = G$, the result will be exactly the same as it is for $s = B$ in (27). This is because, given the assumption of being in an efficient situation ($\tau_G(\phi) =$

$\tau_B(\phi) = 1$), ϕ_g and ϕ_b do not change and, regardless of the environment situation, the difference between probabilities of obtaining a good result and probabilities of obtaining a bad result $(\alpha - \rho)$ is by assumption also the same.

Hence, both in good and bad times the condition for competents to hire experts is the same, being always possible to find some $\Delta^{NS} \geq (w_E - w_N) > 0$ such that $\Delta^{NS} = \Delta^{NS}_G = \Delta^{NS}_B$ and competents delegate to efficient workers for all priors ϕ.

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

Without scapegoating this is always true because,

a) $\delta(1-\lambda) > 0$ since $\delta > 0$ and $\lambda < 1$.

b) $X = (\alpha - \rho)[p(\phi_g) - p(\phi_b)] > 0$ since $\alpha > \rho$ and $p(\phi)$ is monotonically increasing in ϕ ($\frac{\partial p(\phi)}{\partial \phi} = \alpha - \rho > 0$).[20]

c) $V_f > 0$ because $Y_g > 0$ and $Y_b > 0$ since the value function V is monotonically increasing in ϕ as well.[21]

Hence, $\Delta^{NS} > 0$ for all $\phi \in [\lambda\theta, 1 - \lambda(1-\theta)]$. ■

Proof Proposition 6. This proof proceed in two steps. First we solve for the optimal blaming intensities (x , y and z) chosen by superiors and consistent with "consumers" beliefs in equilibrium. Second, considering the results obtained in the blaming stage, we derive conditions for an efficient equilibrium.

Step 1: *Blaming stage equilibrium*

The strategy we follow in this part of the proof is: First, consider as given the "consumers" beliefs about blaming intensities and determine optimal decisions by superiors (both competents (x) and inepts(y))[22] that maximize their utility. 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 the nature. For example, in bad times, when competents hired experts and decide a blaming intensity x, the value function is,

$$V(\phi, x) = p(\phi) - w_E + \delta(1-\lambda)[\alpha V(\phi_g) + (1-\alpha)[xV(\phi_b^{e_c}) + (1-x)V(\phi_b^{s_c})]]$$

where $\phi_b^{e_c}$ and $\phi_b^{s_c}$ are determined by "consumers"'s beliefs about x and y.

[20] More 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 superior's revenues in period t under F is higher than under G.

[21] Following Mailath and Samuelson (2001), let $f_t(\phi, \phi_0, t_0)$ be the distribution of "consumer" 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.

[22] Recall at this point z is not relevant for "consumers" to update beliefs because we are focusing only on efficient equilibria in which competents always hire experts (i.e. $\tau_G(\phi) = \tau_B(\phi) = 1$)

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

$$VD(x) = V(\phi, x') - V(\phi; x) = \delta(1-\lambda)(1-\alpha)(x'-x)[V(\phi_b^{e_c}) - V(\phi_b^{s_c})]$$

1) Assume "consumers" believe $x = y$.

By Corollary 3 and equations (7) and (8), $\phi_b^{e_c} = \phi_b^{s_c} = \phi_b$

Since $V(\phi_b^{e_c}) - V(\phi_b^{s_c}) = 0$, competents are indifferent choosing any $x' \in [0, \overline{x}]$ because, regardless of $(x' - x)$, always $VD(x) = 0$.

Similarly, inepts are indifferent choosing any $y' \in [0, \overline{y}]$ because, regardless of $(y' - y)$, $VD(y) = 0$.

Hence, "consumers" beliefs $x = y$ are correct and consistent with equilibrium strategies.

2) Assume "consumers" believe $x > y$.

By Corollary 3 and equations (7) and (8), $\phi_b^{e_c} > \phi_b > \phi_b^{s_c}$

Since $V(\phi_b^{e_c}) - V(\phi_b^{s_c}) > 0$, competents choose $x' = \overline{x}$ because it maximizes $VD(x)$. Similarly, inepts will choose $y' = \overline{y}$ because it maximizes $VD(y)$.

Only "consumers" beliefs $x = \overline{x}$ and $y = \overline{y}$ will be correct, which are consistent with beliefs $x > y$ solely when $\overline{x} > \overline{y}$.

3) Assume "consumers" believe $x < y$.

By Corollary 3 and equations (7) and (8), $\phi_b^{e_c} < \phi_b < \phi_b^{s_c}$

Since $V(\phi_b^{e_c}) - V(\phi_b^{s_c}) < 0$, competents choose $x' = 0$ because it maximizes $VD(x)$. Similarly, inepts will choose $y' = 0$ because it maximizes $VD(y)$.

Only "consumers" beliefs $x = 0$ and $y = 0$ will be correct, which is not consistent with beliefs in which $x < y$.

Because we are focusing on efficient equilibria, the analysis is done based on beliefs for x and y but a competent type that deviated in the first stage hiring nonexperts will decide any $z \in [0, \overline{z}]$ in case 1, $z = \overline{z}$ in case 2 and $z = 0$ in case 3, exactly for the same arguments explained before for x.

Case 1 supports multiple pooling equilibria in which no further reputation update is delivered from the blaming activity. Case 2 is the only separating equilibrium in which the blaming activity represents an additional reputation updating, but it is only sustained by the case in which $\overline{x} > \overline{y}$. Case 3 is not an equilibrium.

Step 2: *Delegation stage equilibrium*

a) Let $\overline{x} \leq \overline{y}$

Fix ϕ and suppose an efficient situation (i.e., competents always choose to hire experts in both states ($\tau_G(\phi) = \tau_B(\phi) = 1$)).

Consider the case in which **s=B**, competents's value function when hiring experts is,

$$\begin{aligned} V(\phi, DE) \quad &= \quad p(\phi) - w_E + \delta(1-\lambda)[\Pr(g|DE, B)V(\phi_g) + \\ &\Pr(b|DE, B)[\Pr(e_c|b, DE)V(\phi_b^{e_c}) + \Pr(s_c|b, DE)V(\phi_b^{s_c})]] \end{aligned}$$

$$V(\phi, DE) = p(\phi) - w_E + \delta(1-\lambda)[\alpha V(\phi_g) + (1-\alpha)[xV(\phi_b^{e_c}) + (1-x)V(\phi_b^{s_c})]]$$

Since the only possible equilibrium in the blaming stage is a pooling one such that $\phi_b^{e_c} = \phi_b^{s_c} = \phi_b$, we have

$$V(\phi, DE) = p(\phi) - w_E + \delta(1-\lambda)[\alpha V(\phi_g) + (1-\alpha)V(\phi_b)]$$

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

$$V(\phi, DN) = p(\phi) - w_N + \delta(1-\lambda)[\rho V(\phi_g) + (1-\rho)V(\phi_b)]$$

These are exactly the same expressions compared to obtain conditions for an equilibrium without scapegoating in Proposition 5's proof. The same is true when **s=G**. Hence, when $\overline{x} \leq \overline{y}$ the existence of scapegoating does not change efficiency conditions.

b) Let $\overline{x} > \overline{y}$

Even when pooling equilibria in blaming intensities that do not affect conditions for efficiency also exist in this situation, in what follows we discuss the unique separating equilibrium in which $x = \overline{x}$, $y = \overline{y}$ and $z = \overline{z}$ such that $\phi_b^{e_c} > \phi_b > \phi_b^{s_c}$.

Fix ϕ and suppose an efficient situation (i.e. competents always choose to hire experts in both states ($\tau_G(\phi) = \tau_B(\phi) = 1$)). Under results from Corollary 3, given any state $s \in \{B, G\}$, for all feasible ϕ, $\varphi(\varphi(\phi|g)|g) = \phi_{gg} > \phi_g > \phi > \phi_b > \phi_{bb}$ and $\phi_{gi} > \phi_{b_1 i} > \phi_{bi} > \phi_{b_2 i}$ for $i \in \{g, b_1, b_2\}$, calling, just to save notation, the states after a report $b_1 = (b, e_c)$ and $b_2 = (b, s_c)$.

If $s = B$, competent's value function when hiring experts is,

$$\begin{aligned} V(\phi, DE) \quad = \quad & p(\phi) - w_E + \delta(1-\lambda)[\Pr(g|DE, B)V(\phi_g) + \\ & \Pr(b|DE, B)[\Pr(e_c|b, DE)V(\phi_b^{e_c}) + \Pr(s_c|b, DE)V(\phi_b^{s_c})]] \end{aligned}$$

$$V(\phi, DE) = p(\phi) - w_E + \delta(1-\lambda)[\alpha V(\phi_g) + (1-\alpha)[\overline{x}V(\phi_b^{e_c}) + (1-\overline{x})V(\phi_b^{s_c})]]$$

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

$$V(\phi, DN) = p(\phi) - w_N + \delta(1-\lambda)[\rho V(\phi_g) + (1-\rho)[\overline{z}V(\phi_b^{e_c}) + (1-\overline{z})V(\phi_b^{s_c})]]$$

Thus

$$\begin{aligned} V(\phi, DE) - V(\phi; DN) = -(w_E - w_N) + \delta(1-\lambda)[X^B] \\ +\delta^2(1-\lambda)^2\{\Pr(g|DE)Y_g^B + \Pr(b|DE)[\overline{x}Y_{b_1}^B + (1-\overline{x})Y_{b_2}^B]\} \end{aligned}$$

where

$$\begin{aligned} X^B = (\alpha-\rho)p(\phi_g) + (1-\alpha)[\overline{x}p(\phi_b^{e_c}) + (1-\overline{x})p(\phi_b^{s_c})] \\ -(1-\rho)[\overline{z}p(\phi_b^{e_c}) + (1-\overline{z})p(\phi_b^{s_c})] \end{aligned}$$

$$Y_i^B = (\alpha-\rho)V(\phi_{gi}) + (1-\alpha)[\overline{x}V(\phi_{bi}^{e_c}) + (1-\overline{x})V(\phi_{bi}^{s_c})]$$
$$-(1-\rho)[\overline{z}V(\phi_{bi}^{e_c}) + (1-\overline{z})V(\phi_{bi}^{s_c})]$$

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

and, as in the previous proof,

$$\Pr(g|DE) = \alpha + \gamma(1-\rho-\alpha)$$

and

$$\Pr(b|DE) = (1-\alpha) - \gamma(1-\rho-\alpha)$$

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

Just to save notation, rewrite equation (5) $p(\phi) = \Pr(g|C)\phi + Pr(g|I)(1-\phi)$ as $p(\phi) = A\phi + C(1-\phi)$. Then,

$$X^B = (\alpha-\rho)p(\phi_g) + (1-\alpha)[\overline{x}[A\phi_b^{e_c} + C(1-\phi_b^{e_c})] + (1-\overline{x})[A\phi_b^{s_c} + C(1-\phi_b^{s_c})]]$$
$$-(1-\rho)[\overline{z}[A\phi_b^{e_c} + C(1-\phi_b^{e_c})] + (1-\overline{z})[A\phi_b^{s_c} + C(1-\phi_b^{s_c})]]$$

$$X^B = (\alpha-\rho)p(\phi_g) + (1-\alpha)[A[\overline{x}\phi_b^{e_c} + (1-\overline{x})\phi_b^{s_c}] + C[\overline{x}(1-\phi_b^{e_c}) + (1-\overline{x})(1-\phi_b^{s_c})]]$$
$$-(1-\rho)[A[\overline{z}\phi_b^{e_c} + (1-\overline{z})\phi_b^{s_c}] + C[\overline{z}(1-\phi_b^{e_c}) + (1-\overline{z})(1-\phi_b^{s_c})]]$$

Let's define

$$\phi_{b,DE} = \overline{x}\phi_b^{e_c} + (1-\overline{x})\phi_b^{s_c} \tag{28}$$

and

$$\phi_{b,DN} = \overline{z}\phi_b^{e_c} + (1-\overline{z})\phi_b^{s_c} \tag{29}$$

In plain words, these expressions represent the expected reputation competents expect to obtain after a bad result in case of hiring experts ($\phi_{b,DE}$) or nonexperts ($\phi_{b,DN}$)

From the second term of the long expression and from equation (28),

$A\phi_{b,DE} + C(1-\phi_{b,DE}) = p(\phi_{b,DE})$

and from the last term and from equation (29)

$A\phi_{b,DN} + C(1-\phi_{b,DN}) = p(\phi_{b,DN})$

Hence,

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

It's 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 $V(\phi)$).

An equilibrium in which competents only hire experts when $s = B$ requires that $V(\phi, DE) - V(\phi; DN) \geq 0$ for all feasible reputation measures ϕ. A necessary condition for this to happen is that cost differences $(w_E - w_N)$ fulfill

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

where $V_f^B = \Pr(g|DE)Y_g^B + \Pr(b|DE)[\overline{x}Y_{b_1}^B + (1-\overline{x})Y_{b_2}^B]$

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 $\phi \in [\lambda\theta, 1-\lambda(1-\theta)]$

$$\Delta_B^S = \min_{\phi \in [\lambda\theta, 1-\lambda(1-\theta)]} \left\{ \delta(1-\lambda)[X^B + \delta(1-\lambda)V_f^B] \right\} \tag{30}$$

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 blaming. In this case, from tomorrow on it would be possible to have only two possible states $i \in \{g, b\}$. It's straightforward to check that equation (30) is simplified to $V_f^B = \Pr(g|DE)Y_g^B + \Pr(b|DE)Y_b^B$. This expression is used in Proposition 6.

If $s = G$, the proof is identical to the previous one but having $\Pr(g|DE, G) = (1-\rho)$ and $\Pr(g|DN, G) = (1-\alpha)$.

Then,

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

and

$$Y_i^G = (\alpha-\rho)V(\phi_{gi}) + \rho[\overline{x}V(\phi_{bi}^{e_c}) + (1-\overline{x})V(\phi_{bi}^{s_c})] - \alpha[\overline{z}V(\phi_{bi}^{e_c}) + (1-\overline{z})V(\phi_{bi}^{s_c})]$$

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,

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

where $V_f^G = \Pr(g|DE)Y_g^G + \Pr(b|DE)[\overline{x}Y_{b_1}^G + (1-\overline{x})Y_{b_2}^G]$

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 $\phi \in [\lambda\theta, 1-\lambda(1-\theta)]$

$$\Delta_G^S = \min_{\phi \in [\lambda\theta, 1-\lambda(1-\theta)]} \left\{ \delta(1-\lambda)[X^G + \delta(1-\lambda)V_f^G] \right\} \tag{31}$$

As in the previous case, to save notation it is possible to assume a case in which the future does not present scapegoating chances and restrict the states from tomorrow on to just two possibilities, $i \in \{g, b\}$. In this case $V_f^G = \Pr(g|DE)Y_g^G + \Pr(b|DE)Y_b^G$ in equation (31). This is the expression used in Proposition 6. ■

Extension proof Proposition 7. Considering the existence of blaming activities in current and future periods, the only difference arises in the definition of V_f^G as shown in the previous proof, after equation (31)

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

$V_f^G = \Pr(g|DE)Y_g^G + \Pr(b|DE)[\overline{x}Y_{b_1}^G + (1-\overline{x})Y_{b_2}^G]$

Then,

$$V_f - V_f^G = \Pr(g|DE)[Y_g - Y_g^G] + \Pr(b|DE)[Y_b - \overline{x}Y_{b_1}^G - (1-\overline{x})Y_{b_2}^G]$$

This expression will be non-negative whenever
$Y_g - Y_g^G \geq 0$ and $Y_b - \overline{x}Y_{b_1}^G - (1-\overline{x})Y_{b_2}^G \geq 0$

Proof of Proposition 7 in the main text shows that, $Y_g - Y_g^G \geq 0$
Now,
$Y_b - \overline{x}Y_{b_1}^G - (1-\overline{x})Y_{b_2}^G = (\alpha - \rho)[V(\phi_{gb}) - V(\phi_{bb})]$
$-\overline{x}\left[(\alpha-\rho)V(\phi_{gb_1}) + \rho[\overline{x}V(\phi_{bb_1}^{e_c}) + (1-\overline{x})V(\phi_{bb_1}^{s_c})] - \alpha[\overline{z}V(\phi_{bb_1}^{e_c}) + (1-\overline{z})V(\phi_{bb_1}^{s_c})]\right]$
$-(1-\overline{x})\left[(\alpha-\rho)V(\phi_{gb_2}) + \rho[\overline{x}V(\phi_{bb_2}^{e_c}) + (1-\overline{x})V(\phi_{bb_2}^{s_c})] - \alpha[\overline{z}V(\phi_{bb_2}^{e_c}) + (1-\overline{z})V(\phi_{bb_2}^{s_c})]\right]$
with some algebra $Y_b - \overline{x}Y_{b_1}^G - (1-\overline{x})Y_{b_2}^G$ can be decomposed in two sufficient conditions for non negativity.

a)
$(\alpha-\rho)\left[\overline{x}V(\phi_{gb}) + (1-\overline{x})V(\phi_{gb})\right] \geq (\alpha-\rho)\left[\overline{x}V(\phi_{gb_1}) + (1-\overline{x})V(\phi_{gb_2})\right]$
or, which is the same,

$$\overline{x} \leq \frac{V(\phi_{gb}) - V(\phi_{gb_2})}{V(\phi_{gb_1}) - V(\phi_{gb_2})} \leq 1$$

(Recall $\phi_{gb_1} \geq \phi_{gb} \geq \phi_{gb_2}$ from Corollary 3).

b)
$\overline{x}\left[(\alpha\overline{z} - \rho\overline{x})[V(\phi_{bb_1}^{e_c}) - V(\phi_{bb_1}^{s_c})] - (\alpha-\rho)[V(\phi_{bb}) - V(\phi_{bb_1}^{s_c})]\right] +$
$(1-\overline{x})\left[(\alpha\overline{z} - \rho\overline{x})[V(\phi_{bb_2}^{e_c}) - V(\phi_{bb_2}^{s_c})] - (\alpha-\rho)[V(\phi_{bb}) - V(\phi_{bb_2}^{s_c})]\right] \geq 0$
which can be reduced to
$(\alpha\overline{z} - \rho\overline{x})[\overline{x}V(\phi_{bb_1}^{e_c}) + (1-\overline{x})V(\phi_{bb_2}^{e_c})] \geq (\alpha-\rho)V(\phi_{bb})$
The sufficient conditions for this to hold are

$$\overline{x} \geq \frac{V(\phi_{bb}) - V(\phi_{bb_2}^{e_c})}{V(\phi_{bb_1}^{e_c}) - V(\phi_{bb_2}^{e_c})}$$

and the condition already obtained in (16).

$$\overline{z} > z^* = 1 - \frac{\rho}{\alpha}(1-\overline{x})$$

■

Proofs Section 4

Equilibrium with nested activities after successes (similar proof Proposition 6)

Proof. This follows closely the logic and proof's strategies for Proposition 6.

Step 1: *Nested stage equilibrium*

Here we focus on the efficient equilibrium in which competents hire experts.

For example, in bad times, when competents hired experts and decide a boycotting intensity x_g, the value function is,

$$V(\phi, x_g) = p(\phi) - w_E + \delta(1-\lambda)[\alpha[x_g V(\phi_g^s) + (1-x_g)V(\phi_g^f)] + (1-\alpha)V(\phi_b)]$$

where ϕ_g^s and ϕ_g^f are determined by "consumers"'s beliefs about x_g and y_g.

For any deviation from x_g, say to x'_g, we can define,

$$VD(x_g) = V(\phi, x'_g) - V(\phi; x_g) = \delta(1-\lambda)(1-\alpha)(x'_g - x_g)[V(\phi_g^s) - V(\phi_g^f)]$$

1) Assume "consumers" believe $x_g = y_g$.

By equations (18) and (19), $\phi_g^s = \phi_g^f = \phi_g$

Since $V(\phi_g^s) - V(\phi_g^f) = 0$, competents are indifferent choosing any $x'_g \in [0, \overline{x_g}]$ because, regardless of $(x'_g - x_g)$, always $VD(x_g) = 0$. Similarly, inepts are indifferent choosing any $y'_g \in [0, \overline{y_g}]$ because, regardless of $(y'_g - y_g)$, $VD(y_g) = 0$.

Hence, "consumers" beliefs $x_g = y_g$ are correct and consistent with equilibrium strategies.

2) Assume "consumers" believe $x_g > y_g$.

By equations (18) and (19), $\phi_g^s > \phi_g > \phi_g^f$

Competents choose $x'_g = \overline{x_g}$ and inepts $y'_g = \overline{y_g}$. Only "consumers" beliefs $x_g = \overline{x_g}$ and $y_g = \overline{y_g}$ will be correct, which are consistent with beliefs $x_g > y_g$ solely when $\overline{x_g} > \overline{y_g}$. This means superiors in this case choose a nil boycotting intensity.

3) Assume "consumers" believe $x_g < y_g$.

By equations (18) and (19), $\phi_g^s < \phi_g < \phi_g^f$

Competents choose $x'_g = 0$ and inepts $y'_g = 0$. Only "consumers" beliefs $x_g = y_g = 0$ will be correct, not consistent with beliefs in which $x_g < y_g$.

Step 2: *Delegation stage equilibrium*

a) Let $\overline{x_g} \leq \overline{y_g}$

In this situation only a pooling equilibrium exists at the nested stage. Hence, no further updating exists after production and the existence of nested activities after successes does not change efficiency conditions.

b) Let $\overline{x_g} > \overline{y_g}$

Even when pooling equilibria in boycotting intensities that do not affect conditions for efficiency also exist in this situation, in what follows we discuss the unique separating equilibrium in which no boycott occurs, such that $\phi_g^s > \phi_g > \phi_g^f$ since $x_g = \overline{x_g}$ and $y_g = \overline{y_g}$.

Fix ϕ and suppose an efficient situation (i.e. competents always choose to hire experts in both states ($\tau_G(\phi) = \tau_B(\phi) = 1$)).

If $\boldsymbol{s = B}$, competent's value function when hiring experts is,

$$V(\phi, DE) = p(\phi) - w_E + \delta(1-\lambda)[\alpha[\overline{x_g}V(\phi_g^s) + (1-\overline{x_g})V(\phi_g^f)] + (1-\alpha)V(\phi_b)]$$

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

$$V(\phi, DN) = p(\phi) - w_N + \delta(1-\lambda)[\rho[\overline{y_g}V(\phi_g^s) + (1-\overline{y_g})V(\phi_g^f)] + (1-\rho)V(\phi_b)]$$

Then

$$V(\phi, DE) - V(\phi; DN) = -(w_E - w_N) + \delta(1-\lambda)[X^{BS}]$$
$$+\delta^2(1-\lambda)^2\{\Pr(g|DE)[\overline{x_g}Y_{g_1}^{BS} + (1-\overline{x_g})Y_{g_2}^{BS}] + \Pr(b|DE)Y_b^{BS}\}$$

where

$$X^{BS} = \alpha p(\phi_{g,DE}) - \rho p(\phi_{g,DN}) - (\alpha - \rho)p(\phi_b)$$

with

$$\phi_{g,DE} = \overline{x_g}\phi_g^s + (1-\overline{x_g})\phi_g^f$$
$$\phi_{g,DN} = \overline{y_g}\phi_g^s + (1-\overline{y_g})\phi_g^f$$

and

$$Y_i^{BS} = \alpha[\overline{x_g}V(\phi_{gi}^s) + (1-\overline{x_g})V(\phi_{gi}^f)] - \rho[\overline{y_g}V(\phi_{gi}^s) + (1-\overline{y_g})V(\phi_{gi}^f)] - (\alpha-\rho)V(\phi_{bi})$$

for $i \in \{g_1, g_2, b\}$. (To save notation $g_1 = (g, s)$ and $g_2 = (g, f)$).

and, as in previous proofs,

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

A necessary condition for an efficient equilibrium is that cost differences $(w_E - w_N)$ fulfill

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

where $V_f^{BS} = \Pr(g|DE)[\overline{x_g}Y_{g_1}^{BS} + (1-\overline{x_g})Y_{g_2}^{BS}] + \Pr(b|DE)Y_b^{BS}$

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

$$\Delta_B^{SS} = \min_{\phi\in[\lambda\theta,1-\lambda(1-\theta)]} \left\{\delta(1-\lambda)[X^{BS} + \delta(1-\lambda)V_f^{BS}]\right\} \tag{32}$$

To save notation it is possible to assume a case in which the future does not present nested activities. In this case from tomorrow on it would be possible to have only two states $i \in \{g, b\}$. It's straightforward to check that equation (32) is simplified to $V_f^{BS} = \Pr(g|DE)Y_g^{BS} + \Pr(b|DE)Y_b^{BS}$.

If $s = G$, the proof is identical to the previous one, but having $\Pr(g|DE, G) = (1-\rho)$ and $\Pr(g|DN, G) = (1-\alpha)$.

Then,

$$X^{GS} = (1-\rho)p(\phi_{g,DE}) - (1-\alpha)p(\phi_{g,DN}) - (\alpha-\rho)p(\phi_b)$$

and

$$Y_i^{GS} = (1-\rho)[\overline{x_g}V(\phi_{gi}^s) + (1-\overline{x_g})V(\phi_{gi}^f)] - (1-\alpha)[\overline{y_g}V(\phi_{gi}^s) + (1-\overline{y_g})V(\phi_{gi}^f)] - (\alpha-\rho)V(\phi_{bi})$$

for $i \in \{g_1, g_2, b\}$

Hence the condition for competents to hire experts and to achieve the efficient situation is,

$$(w_E - w_N) \leq \Delta_G^{SS} = \min_{\phi\in[\lambda\theta,1-\lambda(1-\theta)]} \left\{\delta(1-\lambda)[X^{GS} + \delta(1-\lambda)V_f^{GS}]\right\} \tag{33}$$

where $V_f^{GS} = \Pr(g|DE)[\overline{x_g}Y_{g_1}^{GS} + (1-\overline{x_g})Y_{g_2}^{GS}] + \Pr(b|DE)Y_b^{GS}$ or $V_f^{GS} = \Pr(g|DE)Y_g^{GS} + \Pr(b|DE)Y_b^{GS}$ in equation (33) if in the future only states $i \in \{g, b\}$ are possible. ■

Efficiency of nested activities after successes (similar proof Proposition 7)

Proof. This follows closely the logic in the proof of Proposition 7. Here we will show nested activities after successes promote efficiency, as opposed to the case with scapegoating (nested activity after failures).

As shown in the text by equations (20)-(23), $\phi_{g,DE} > \phi_g > \phi_{g,DN} > \phi_b$.

We have to show $\Delta_B^{SS} \geq \Delta^{NS}$, because we will show later that $\Delta_G^{SS} \geq \Delta_B^{SS}$. It's enough to prove the following two claims.

Claim 1) $X^{BS} \geq X$ for all ϕ.

$$X^{BS} - X = \alpha[p(\phi_{g,DE}) - p(\phi_g)] - \rho[p(\phi_{g,DN}) - p(\phi_g)]$$

which is non-negative since $\alpha > \rho$ by assumption and $\phi_{g,DE} \geq \phi_{g,DN}$.

Claim 2) $V_f^{BS} \geq V_f$ for all ϕ.

$$V_f^{BS} - V_f = \Pr(g|DE)[Y_g^B - Y_g] + \Pr(b|DE)[Y_b^B - Y_b]$$

and

$$Y_i^{BS} - Y_i = (\alpha\overline{x_g} - \rho\overline{y_g})[V(\phi_{gi}^s) - V(\phi_{gi}^f)] - (\alpha - \rho)[V(\phi_{gi}) - V(\phi_{gi}^f)]$$

for $i \in \{g, b\}$. It is non-negative when $(\alpha\overline{x_g} - \rho\overline{y_g}) \geq (\alpha - \rho)$ since $\phi_{gi}^s \geq \phi_{gi}$. Then the sufficient condition for non negativity is

$$\overline{x_g} \geq x_g^* = 1 - \frac{\rho}{\alpha}(1 - \overline{y_g})$$

■

Machiavellian Effect with nested activities after successes (similar proof Proposition 10)

Proof. This follows closely the logic in the proof of Proposition 10.

As shown in the text by equations (20)-(23), $\phi_{g,DE} > \phi_g > \phi_{g,DN} > \phi_b$.

We have to show $\Delta_G^{SS} \geq \Delta_B^{SS}$ for a "Machiavellian Effect" to exist. It's enough to prove the following two claims.

Claim 1) $X^{GS} \geq X^{BS}$ *for all* ϕ.

$$X^{GS} - X^{BS} = (1 - \alpha - \rho)[p(\phi_{g,DE}) - p(\phi_{g,DN})]$$

which is non-negative since $\alpha + \rho < 1$ by assumption and $\phi_{g,DE} \geq \phi_{g,DN}$

Claim 2) $V_f^{GS} \geq V_f^{BS}$ *for all* ϕ

$$V_f^{GS} - V_f^{BS} = \Pr(g|DE)[Y_g^{GS} - Y_g^{BS}] + \Pr(b|DE)[Y_b^{GS} - Y_b^{BS}]$$

which is non-negative because
$Y_i^{GS} - Y_i^{BS} = (1 - \alpha - \rho)(\overline{x_g} - \overline{y_g})[V(\phi_{gi}^s) - V(\phi_{gi}^f)] \geq 0$
for $i \in \{g, b\}$ since $\alpha + \rho < 1$ by assumption and $\phi_{gi}^s \geq \phi_{gi}^f$. ■