

# Reputation from Nested Activities

Guillermo L. Ordoñez\*

August, 2011

---

\*Department of Economics, Yale University. Address: 28 Hillhouse Av., Rm 208, New Haven, CT, 06511. E-mail: guillermo.ordonez@yale.edu. Phone: (203) 432-8320. I would like to thank the editor, an anonymous referee, Christian Hellwig, David K. Levine, Joe Ostroy, Larry Samuelson, Vasiliki Skreta, Jeroen Swinkels, Bill Zame, and seminar participants at UCLA, Minnesota, and Washington University in St. Louis, for useful comments. The usual waiver of liability applies.

### **Abstract**

Principals usually try to elicit the quality and behavior of agents from their performance. While sometimes the success or failure in production do not provide accurate signals about the agents, there may be activities not directly related to production that constitute a more precise source of information about them. I show that, when agents face reputation concerns, introducing these activities after a success improves efficiency, while introducing them after a failure reduces efficiency. Hence, nesting activities in the right way may offer a cheap toolbox to provide incentives. As an illustration, I consider a model where reputation concerns drive the hiring decisions of managers in a firm and I show how scapegoating, an activity "nested" after failures in production, generates inefficiencies. While hiring efficient workers increases the probability of success, hiring less efficient workers provides a buffer against reputation losses from failures, since managers can blame them more easily.

Keywords: nested reputation, delegation, efficiency, scapegoating.

JEL Codes: D81, D21, D73, M51

# 1 Introduction

In many firms and organizations, the success or failure in production may constitute a poor source of information about the employees' quality and efforts. Usually, many different factors jointly determine the probability of success, complicating the inference about the performance of a single factor. When agents have reputation concerns, their incentives to put effort and invest in quality critically depend on how well the principal can learn from agents' performance. If learning is difficult, there may not be enough incentives for agents to behave and exert efforts.

However, it is usually possible to find a wide array of activities, not directly related to production, that provide cleaner and more accurate signals about agents' quality and efforts. In this paper I argue that nesting these activities after a successful production increases the probability of efficient effort provision, while nesting them after failures in production reduces that probability. Even when these activities do not affect production directly, they do so indirectly through the decisions governing the probability of success in production. This result is important because nesting activities in the right way may constitute a cheap tool to provide incentives.

To illustrate this point I explore a particular application to hiring decisions by reputation-concerned managers in firms that provide a service and study the effects of scapegoating on efficient hiring. I interpret scapegoating as an activity that managers can use to explain failures, which even when irrelevant, constitutes an additional signal of their competence.

Even when used just as an example in this paper, scapegoating is itself an interesting phenomenon. A quick review of major newspapers in many countries shows that people commonly condemn scapegoating behavior. This attitude stems not only from its unfairness but also from its negative effects on efficiency and performance in organizations. In fact, this general view has been widely used to justify recent institutional reforms designed to improve efficiency by reducing scapegoating possibilities.<sup>1</sup> So far, however, no model has been developed to formalize this conventional and seemingly well-accepted idea.

I introduce scapegoating into an environment with hidden information about managers' types and hiring decisions. In the model, managers need to hire a worker to provide a service, which can turn out being a good or a bad service. Managers can be competent or inept while workers can be experts or nonexperts. Inept managers' hiring decisions do not affect performance, since they are always the sole responsible of providing a bad service. Contrarily,

---

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

competent managers are never responsible for a bad service and hiring experts reduces the probability of such an occurrence. Since competent managers are indeed not responsible for a bad service, they are more effective than inept managers in blaming workers if something goes wrong. Hence, both production and scapegoating results are useful elements for consumers to learn about the manager's type. When making their hiring decisions, inept managers always hire nonexperts, while competent managers know that hiring experts increases the probability of providing a good service but also that hiring nonexperts constitutes a better buffer against reputation losses in case of providing a bad service.

I assume hiring experts is efficient and I derive conditions for efficiency to be sustained as an equilibrium, with and without scapegoating. I show it is more difficult, and sometimes even impossible, to achieve efficiency as an equilibrium with scapegoating. This is because the efficient hiring of experts hinders the use of blaming to maintain reputation after failures (since it is harder to blame experts than nonexperts), reducing the expected reputation gains from working with experts.

I extend the paper in two directions. First I show that, contrarily to scapegoating, activities used exclusively after successes increase the chances of an efficient equilibrium, since they exploit reputation forces in the right direction, without requiring monetary resources or costly incentives. Second, I show that considering reputation from nested activities allows us to identify a "Machiavellian effect,"<sup>2</sup> which can be restated as "Managers tend to hire nonexperts during bad times and experts during good times." The intuition is simple. In good times the probability of success is higher and managers rely less on scapegoating, inducing them to take more efficient hiring decisions. This result may be potentially relevant in explaining lengthy recoveries from recessions, characterized by less efficient hiring.

The literature on reputation is very large.<sup>3</sup> However, to the best of my knowledge, nested reputation models do not exist. Although this paper shares some features with the literature on "reputation spillover" (as in Cole and Kehoe (1996)) and "analogy-based learning" (as in Miettinen (2010)), the logic is not the same. While spillovers deal with multi-dimensional types and analogies deal with grouping populations of differing characteristics, in my work reputation is constructed around a single hidden type, but through several nested stages and steps. This paper is then a first step towards understanding which is the optimal sequence of activities that induce efficient actions by fully exploiting reputation concerns.

---

<sup>2</sup>In his famous book *The Prince*, Machiavelli wrote, "Princes should delegate to others the enactment of unpopular measures and keep in their own hands the distribution of favours." Machiavelli's argument was that princes should delegate when the probability of having a success is low and work by themselves if it is high. In this way, princes would be able to blame others if something goes wrong, thereby maintaining their reputation. More recently, Alesina and Tabellini (2005, 2007, 2008) formally modeled this behavior among politicians.

<sup>3</sup>Starting with Kreps and Wilson (1982) and Milgrom and Roberts (1982), important contributions on reputation models are Fudenberg and Levine (1989, 1992), Mailath and Samuelson (1998, 2001, 2006), Ely and Valimaki (2003), Tadelis (2003), Cripps, Mailath, and Samuelson (2004), and Levy (2005).

Even when just an application of reputation from nested activities, this paper also contributes to an almost nonexistent literature on scapegoating. Despite the recognition of strategic reasons for scapegoating in the social-psychology literature (Bell and Tetlock (1989); Douglas (1995)), formal economic studies of this behavior are new and sparse. Dezsó (2004) analyzes the conditions under which random firing of potential innocents (scapegoats) is a reaction to failures in order to maintain reputation. He focuses on firing and not on hiring, without being able to analyze the impact on efficiency. Segendorff (2000) analyzes the possible hiring of scapegoats using a signaling game, also without analyzing efficiency consequences. Winter (2001) finds that, under some circumstances, in order to provide better incentives to top levels in an organization, it may be optimal for middle levels to bear more responsibility, an aspect he labels "scapegoating." Again, he does not consider reputation effects or hiring decisions.<sup>4</sup>

Finally, my work also contributes to the literature that studies nonstandard ways to induce workers and managers to provide effort. Alonso-Pauli and Perez-Castrillo (2010) discuss the use of Codes of Best Practice to improve manager control, which depends on the speed of learning about those managers. Cappelen and Tungodden (2009) study the use of egalitarian rewards across workers to provide incentives. Koch and Peyrache (2008) show that creating ambiguity about an agent's performance can generate career concerns in the absence of ex ante uncertainty about an agent's type, hence introducing incentives for effort exertion. Mylov and Schmitz (2008) study the optimal assignment of workers to tasks over time in order to maximize the incentives to provide efforts.

In Section 2, I discuss the basic model of reputation with hiring and scapegoating, a special application of the proposed "nested reputation" environment. In Section 3, I analyze the conditions for the existence of an efficient equilibrium and show how and when scapegoating leads to inefficiency and to the "Machiavellian effect." In Section 4, I explore the efficacy of "nested" activities after successes to induce efficiency by exploiting reputation concerns, without relying on the use of costly incentives. In Section 5, I make some final remarks.

## 2 The Model

### 2.1 Description

Consider a firm's manager who is responsible for providing a service to consumers. To perform his activity, the manager has to hire a worker and delegate part of the job. The match of

---

<sup>4</sup>Empirical studies about scapegoating are even less common. An exception is Huson, Malatesta, and Parrino (2004), who developed a moral hazard driven scapegoat hypothesis based on agency models to study the impact of managerial succession on firm performance.

a manager and a worker produces a single non-deterministic outcome, which can be a good service ( $g$ ) or a bad service ( $b$ ).

There are two types of managers and two types of workers. Managers can be competent ( $C$ ) or inept ( $I$ ). Workers can be experts ( $E$ ) or nonexperts ( $N$ ). The combination of these types generates four possible matches that differ in their probability of providing a good service and in the responsibility of each agent for providing a bad service. The production technology is as follows: A match with an inept manager provides a good service with probability  $\beta$ , independent of the worker's type, and the manager is the only responsible for the provision of a bad service. A match of a competent manager and a nonexpert also provides a good service with probability  $\beta$ , but the worker is the only responsible for a bad service. Finally, a match of a competent manager and an expert provides a good service with probability  $\alpha > \beta$ , and the worker is also the only responsible in case of a bad service.<sup>5</sup>

If the firm provides a bad service, managers can show some costless evidence (which may or may not be accurate) trying to blame the worker. I define this activity as scapegoating, a nested second stage that only occurs after a bad service, not after a good one.<sup>6</sup> The success ( $s$ ) or failure ( $f$ ) of scapegoating to blame the worker is non-deterministic and depends on the match combination. The scapegoating technology is as follows: The probability of success from scapegoating is  $\eta$  if the manager is inept,  $\gamma$  if the manager is competent and the worker is an expert and  $\rho$  if the manager is competent and the worker is a nonexpert. Given the assignment of responsibilities described above, it is natural to assume that  $\eta \leq \gamma \leq \rho \leq 1$ .

I make three assumptions just for convenience on notation. First, there is no punishment from blaming results, other than potential reputation loses. Second, in case of no scapegoating, consumers always blame the manager. Third, when consumers blame the manager, they do not know whether it was a failure of scapegoating or a lack of scapegoating.

After detailing the production and scapegoating technologies, now I describe preferences. Consumers repeatedly receive the service provided by firms under managers' commands.<sup>7</sup> This generates two possible utility levels in each period: 1 if the service is good ( $u(g) = 1$ ) and 0 if it is bad ( $u(b) = 0$ ). Consumers do not get any utility from scapegoating or from knowing who was the responsible of a bad service, other than the potential learning about the manager's type.

---

<sup>5</sup>The assumption of a same stochastic process for inept managers and competent managers working with nonexperts is introduced just for convenience (we can work with just two parameters  $\alpha$  and  $\beta$ , rather than four) and do not change the main results.

<sup>6</sup>This assumption is key to analyze the incentives from activities that are restricted to occur only after certain results, which we interpret as nested activities. Later we discuss the conditions under which this assumption is an equilibrium result.

<sup>7</sup>There is a continuum of identical consumers of unit mass (such that no single individual can affect the future play of the game) who buy the service from the manager.

With respect to the information structure, I assume consumers do not observe the managers' type and hiring decisions. Consumers only observe whether the service was good or bad and whether the scapegoating was successful or not. Using this information, they update the managers' reputation, which is defined as the probability that the manager is competent,  $\Pr(C) = \phi$ . Since consumers buy the service before production takes place, they pay the expected utility and not the realized utility from the service. Managers with a better reputation tend to provide good services more likely and then consumers are willing to pay a higher price for that service. This is why managers are so concerned about reputation, while consumers are concerned only about the expected utility derived from the service.

I also assume that managers observe workers' types when hiring. Nonexperts charge a wage normalized to 0 and experts an exogenous wage  $w < (\alpha - \beta)$ . This inequality implies it is always efficient that competent managers hire experts. In other words, if consumers knew managers' types and hiring decisions, they would be willing to pay a premium for competent managers to always hire experts. Finally, I assume a manager can be replaced at the end of each period with a fixed probability  $\lambda$  by another manager who is competent with probability  $\phi_0 \in (0, 1)$ . This substitution is non-observable.<sup>8</sup>

It is important to highlight at this point that both workers and consumers are virtually "machines", who do not have any interesting role other than providing a compelling environment to study managers' decisions. Workers can be replaced by any other input with non-observable types and differential costs. Consumers just update reputation and pay a price accordingly. This mechanical view of workers and consumers simplify the exposition to focus on the interaction between the hiring and the scapegoating decisions of reputation concerned managers.

## 2.2 Timing

The timing of the model is as follows:

0) A manager receives the payment for the period, before production takes place. The payment is not contingent on the service quality, just on the manager's reputation  $\phi$  at the beginning of the period.

1) Each manager hires experts or nonexperts. Consumers do not observe this decision.

2) The outcome of production is realized. All agents in the economy observe whether the service was good ( $g$ ) or bad ( $b$ ).

---

<sup>8</sup>This is just a technical assumption needed to sustain an efficient equilibrium in the long run, as discussed in Mailath and Samuelson (2001) and Cripps, Mailath, and Samuelson (2007).

3) The manager can blame the worker for a bad service. This scapegoating action can succeed or fail.

4) With probability  $\lambda$  the manager is replaced by another one, who is competent with a probability  $\phi_0$ .

### 2.3 Equilibrium Definition

First, we restrict behavior to be Markov in order to eliminate equilibria that depend on implausible degrees of coordination between the manager's behavior and consumers beliefs about that manager's behavior. Under uncertainty about the manager's type, the state variable is the probability assigned by consumers to the manager being competent (i.e., the reputation denoted as  $\phi$ ).

Before production, a Markov hiring strategy for competent managers is a mapping  $\tau^i : [0, 1] \rightarrow [0, 1]$ , where  $\tau^i(\phi)$  is the probability a manager  $i \in \{C, I\}$  hires an expert when reputation is  $\phi \in [0, 1]$ . After providing a bad service, a Markov scapegoating strategy is a mapping  $\nu_j^i : [0, 1] \rightarrow [0, 1]$ , where  $\nu_j^i(\phi)$  is the probability a manager  $i \in \{C, I\}$  who hired a worker  $j \in \{E, N\}$  blames the worker.

Consumers are described by a Markov belief function  $p : [0, 1] \rightarrow [0, 1]$ , where  $p(\phi)$  is the probability that consumers assign to receiving a good outcome, given a reputation  $\phi \in [0, 1]$ . Since utilities have been normalized to 1 if the service is good and 0 if it is bad, then  $p(\phi)$  is also the expected utility buying the service from a manager with reputation  $\phi$ .

In a Markov perfect equilibrium, managers maximize expected discounted profits, consumers beliefs about managers' actions are correct and they use Bayes' rule when possible to update beliefs about managers' types. There are two possible rounds of updating: the update after the service has been provided, ( $\Pr(C|g)$  and  $\Pr(C|b)$ ) and the potential update after scapegoating results ( $\Pr(C|b, s)$  and  $\Pr(C|b, f)$ ).

The value function for a manager  $i \in \{C, I\}$  is

$$V^i(\phi) = \max_{\tau^i, \nu_j^i} \{p(\phi) - \tau^i w + \delta(1 - \lambda)E[V(\phi')|\tau^i, \nu_j^i]\}, \quad (1)$$

where  $\delta \in [0, 1]$  is the discount factor, and expectations are constructed over possible reputation levels next period,  $\phi'(\tau^i, \nu_j^i)$ .

**Definition 1** *A Markov perfect equilibrium is: probabilities of hiring experts  $\tau^i(\phi)$ , probabilities of scapegoating  $\nu_j^i(\phi)$ , consumers' expected utility  $p(\phi)$  and posterior beliefs  $\varphi = \Pr(C|R, \phi)$  (where  $R$  are the three possible results  $R \in \{g; (b, s); (b, f)\}$ ), such that*

1. *Hiring:  $\tau^i(\phi)$  maximize the value functions  $V^i(\phi)$  (equation 1) for  $i \in \{C, I\}$  and all possible reputations  $\phi$ .*
2. *Scapegoating:  $\nu_j^i(\phi)$  maximize the value functions  $V^i(\phi)$  (equation 1) for  $i \in \{C, I\}$ ,  $j \in \{E, N\}$  and all possible reputations  $\phi$ .*
3. *Expected utility of consumers: Probability consumers assign to receiving a good service from a manager with reputation  $\phi$ .*

$$p(\phi) = \Pr(g|\phi) = \Pr(g|C)\phi + \Pr(g|I)(1 - \phi) \quad (2)$$

4. *Beliefs about competence (updated using Bayes' rule).<sup>9</sup>*

(a) *Update after a good service ( $g$ )*

$$\varphi(\phi|g) \equiv \phi_g = (1 - \lambda) \Pr(C|g) + \lambda\phi_0 \quad (3)$$

(b) *Update after a successful scapegoating ( $b, s$ )*

$$\varphi(\phi|b, s) \equiv \phi_b^s = (1 - \lambda) \Pr(C|b, s) + \lambda\phi_0 \quad (4)$$

(c) *Update after a failing scapegoating ( $b, f$ )*

$$\varphi(\phi|b, f) \equiv \phi_b^f = (1 - \lambda) \Pr(C|b, f) + \lambda\phi_0 \quad (5)$$

5. *Consumers' beliefs about managers' actions are correct.*

### 3 Efficient Equilibrium, Inefficient Scapegoating

A fundamental question in this paper is: Does scapegoating really reduce the probability of achieving an efficient outcome? Hence, we focus on the conditions for an efficient outcome to be sustained as an equilibrium.<sup>10</sup> Recall our assumption  $0 < w < (\alpha - \beta)$ , implies that efficiency is achieved when competent managers hire experts, regardless of their reputation.

A first characterization of equilibrium is that inept managers never hire experts. This is clear since, for inept managers, hiring experts does not translate into higher probabilities of success

<sup>9</sup>Explicit expressions as functions of strategies are provided in the proof of Proposition 2 below.

<sup>10</sup>This model has multiple equilibria, including a very inefficient one that may arise without conditions, in which managers only hire nonexperts. Intuitively, in this case, consumers do not have statistical elements to update their beliefs about the manager's type. Since there are no gains in terms of reputation from hiring experts, competent managers optimally prefer to hire nonexperts rather than experts and pay higher wages.

in producing or blaming, but implies they have to pay positive wages. Then it is a dominant strategy for inept managers to hire nonexperts (i.e.,  $\tau^I(\phi) = 0$ , for all feasible  $\phi$ ). Given this result, we need to focus on competent managers' hiring decisions. In what follows, for simplicity, I just get rid of the superscript  $C$  when referring to competent managers' hiring strategies and value functions.

I show the sufficient condition for efficiency to be sustained as an equilibrium (i.e.,  $\tau^C(\phi) = 1$ , for all feasible  $\phi$ ) is  $w < \Delta^S$ , where  $\Delta^S$  is a cutoff value we will characterize when scapegoating is allowed. As a benchmark we will also characterize  $\Delta^{NS}$ , the cutoff when scapegoating is not allowed. Whenever  $\gamma > \eta$  and scapegoating abilities are high enough (specifically when a sufficient condition  $\rho \geq 1 - \frac{(1-\alpha)}{(1-\beta)}(1-\gamma)$  holds),

$$\Delta^{NS} \geq \Delta^S \quad \text{and} \quad \Delta^{NS} > 0. \quad (6)$$

Given wages in the economy, the first inequality, which is in fact typically strict, summarizes the main conclusion of the paper: Scapegoating makes the condition for an efficient equilibrium more difficult to hold. The second inequality means that in the absence of scapegoating it is always possible to find a positive wage that sustains efficiency, which is not necessarily the case with scapegoating.

### 3.1 Conditions for Efficient Equilibrium

As a first step, I derive the sufficient condition for the existence of an efficient equilibrium without scapegoating. In this case, only two possible states ( $g$  and  $b$ ) are possible, since there is no blaming activity allowed after the provision of a bad service (*nobody asks why things went wrong!*). Recall the reputation after a bad service is  $\phi_b$ . Similarly, I denote the reputation after two consecutive bad services ( $\varphi(\varphi(\phi|b)|b)$ ) as  $\phi_{bb}$ . The proof is in the Appendix.

#### **Proposition 1** *Efficient Equilibrium without Scapegoating*

Suppose  $\lambda \in (0, 1)$ ,  $\phi \in [\lambda\phi_0, 1 - \lambda(1 - \phi_0)]$ ,  $\delta \in (0, 1)$ , and  $\phi_0 \in (0, 1)$ . If scapegoating is not allowed (no "blaming" stage), there exists a positive cutoff

$$\Delta^{NS} = \min_{\phi \in [\lambda\phi_0, 1 - \lambda(1 - \phi_0)]} \{ \delta(1 - \lambda)[X^{NS} + \delta(1 - \lambda)Y^{NS}] \} > 0 \quad (7)$$

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

where

$$\begin{aligned} X^{NS} &= (\alpha - \beta)[p(\phi_g) - p(\phi_b)] \\ Y^{NS} &= \alpha Y_g^{NS} + (1 - \alpha) Y_b^{NS} \end{aligned}$$

and  $Y_k^{NS} = (\alpha - \beta)(V(\phi_{gk}) - V(\phi_{bk}))$  for  $k \in \{g, b\}$  and  $V(\phi)$  defined in equation (1).

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

**Proposition 2 Efficient Equilibrium with Scapegoating**

Suppose  $\lambda \in (0, 1)$ ,  $\phi \in [\lambda\phi_0, 1 - \lambda(1 - \phi_0)]$ ,  $\delta \in (0, 1)$ , and  $\phi_0 \in (0, 1)$ . If scapegoating is allowed,

a) If  $\gamma \leq \eta$ ,  $\nu_j^I(\phi) = \nu_E^C(\phi) \frac{\gamma}{\eta}$  for all  $\phi$  and all  $\nu_E^C(\phi) \in [0, 1]$ , then conditions for an efficient equilibrium are exactly the same as the case without scapegoating in Proposition 1.

b) If  $\gamma > \eta$ ,  $\nu_j^I(\phi) = 1$  and  $\nu_j^C(\phi) = 1$  for  $j \in \{E, N\}$  and all  $\phi$ , then there exists a, not necessarily positive, cutoff

$$\Delta^S = \min_{\phi \in [\lambda\phi_0, 1 - \lambda(1 - \phi_0)]} \{ \delta(1 - \lambda)[X^S + \delta(1 - \lambda)Y^S] \} \quad (8)$$

such that, for all wages  $0 < w \leq \Delta^S$ , the efficient pure strategy profile in which competent managers always hire experts is a Markov perfect equilibrium,

where

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

with

$$\phi_{b,E} = \gamma\phi_b^s + (1 - \gamma)\phi_b^f \quad (9)$$

$$\phi_{b,N} = \rho\phi_b^s + (1 - \rho)\phi_b^f \quad (10)$$

and

$$Y^S = \alpha Y_g^S + (1 - \alpha) Y_b^S$$

such that, for  $k \in \{g, b\}$ ,

$$Y_k^S = (\alpha - \beta)V(\phi_{gk}) + (1 - \alpha)[\gamma V(\phi_{bk}^s) + (1 - \gamma)V(\phi_{bk}^f)] - (1 - \beta)[\rho V(\phi_{bk}^s) + (1 - \rho)V(\phi_{bk}^f)]$$

Before going to the main proposition of the paper, it is important to emphasize two features of the solution. First, the non-scapegoating case is just a particular case of the scapegoating

solution. If  $\gamma \leq \eta$ , both cases are in fact exactly the same. If  $\gamma > \eta$ , as  $\eta, \gamma, \rho \rightarrow 0$  (maintaining the relation  $\rho \geq \gamma > \eta$ ) always  $\phi_{b,N} \rightarrow \phi_{b,E} \rightarrow \phi_b$  (as can be checked easily from equations (4), (5), (9) and (10)). In this case  $X^S \rightarrow X^{NS}$ , and  $Y^S \rightarrow Y^{NS}$ , which implies that cutoffs approach ( $\Delta^S \rightarrow \Delta^{NS}$ ) as blaming loses effectiveness, or that Proposition 2 approaches Proposition 1.

Second, in the more natural situation in which  $\gamma > \eta$ ,  $\Delta^S$  can be negative. If this is the case, no positive wage can possibly support an efficient equilibrium, which is never the case without scapegoating. In the remainder of the paper, and unless stated otherwise, when referring to the scapegoating case, I refer specifically to the case in which blaming is naturally easier for competent managers, who are indeed not responsible for the failure (this is,  $\gamma > \eta$ ).

### 3.2 Scapegoating Inefficiency

Here, I show the difficulties scapegoating imposes in achieving efficiency. I do this by proving that the conditions for efficiency with scapegoating (when  $\gamma > \eta$  from Proposition 2) are more difficult to hold than the conditions for efficiency without scapegoating (from Proposition 1).

#### Proposition 3 *Scapegoating Inefficiency*

*Suppose  $\lambda \in (0, 1)$ ,  $\phi \in [\lambda\phi_0, 1 - \lambda(1 - \phi_0)]$ ,  $\delta \in (0, 1)$ ,  $\phi_0 \in (0, 1)$  and competent managers blame better than inept managers ( $\gamma > \eta$ ). It is always possible to find a  $\rho \geq \rho^* = 1 - \frac{(1-\alpha)}{(1-\beta)}(1 - \gamma)$  such that the range of wages  $w > 0$  that supports an efficient situation is smaller with scapegoating than without it.*

**Proof.** We need to prove that  $\Delta^{NS} \geq \Delta^S$  for all  $\phi \in (0, 1)$ . This proof is based on the simpler case in which scapegoating is possible only in the current period, not in the future. The conclusion for the more general case is the same, but it is characterized by more complex statements (shown in the Appendix). I consider only the relevant case in which  $\gamma > \eta$  and there is a separating blaming equilibrium such that  $\phi_b^s > \phi_b > \phi_b^f$ .

I proceed in three steps. First, I show that  $\phi_{b,N} \geq \phi_{b,E}$ , second, that  $\phi_{b,E} \geq \phi_b$  (as defined in Proposition 2), and finally, that  $\Delta^{NS} \geq \Delta^S$  by proving that  $X^{NS} + \delta(1 - \lambda)Y^{NS} \geq X^S + \delta(1 - \lambda)Y^S$  for all feasible  $\phi$ .

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

Considering beliefs about hiring in the efficient equilibrium (i.e.,  $\tau(\phi) = 1$ ), we can define

$$\hat{\phi}_b = Pr(C|b) = \frac{(1 - \alpha)\phi}{(1 - \alpha)\phi + (1 - \beta)(1 - \phi)} < \phi$$

Recall  $\phi_b = (1 - \lambda)\hat{\phi}_b + \lambda\phi_0$ , hence  $\hat{\phi}_b$  represents the standard Bayes updating after a bad outcome and before any blaming activity, which is not adjusted by  $\lambda$  because it happens before the period ends and a replacement occurs.

From equations (10), (4), and (5).

$$\begin{aligned}\phi_{b,N} &= \rho\phi_b^s + (1 - \rho)\phi_b^f \\ \phi_{b,N} &= (1 - \lambda) \left[ \rho \frac{\gamma\hat{\phi}_b}{\gamma\hat{\phi}_b + \eta(1 - \hat{\phi}_b)} + (1 - \rho) \frac{(1 - \gamma)\hat{\phi}_b}{(1 - \gamma)\hat{\phi}_b + (1 - \eta)(1 - \hat{\phi}_b)} \right] + \lambda\phi_0\end{aligned}$$

and, from equations (9), (4), and (5),

$$\begin{aligned}\phi_{b,E} &= \gamma\phi_b^s + (1 - \gamma)\phi_b^f \\ \phi_{b,E} &= (1 - \lambda) \left[ \gamma \frac{\gamma\hat{\phi}_b}{\gamma\hat{\phi}_b + \eta(1 - \hat{\phi}_b)} + (1 - \gamma) \frac{(1 - \gamma)\hat{\phi}_b}{(1 - \gamma)\hat{\phi}_b + (1 - \eta)(1 - \hat{\phi}_b)} \right] + \lambda\phi_0,\end{aligned}$$

Subtracting both expressions:

$$\phi_{b,N} - \phi_{b,E} = \frac{(1 - \lambda)\hat{\phi}_b(1 - \hat{\phi}_b)(\rho - \gamma)(\gamma - \eta)}{\eta(1 - \eta) + \hat{\phi}_b(\gamma - \eta)[1 - 2\eta - \hat{\phi}_b(\gamma - \eta)]} \geq 0, \quad (11)$$

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

Subtract  $\phi_b = (1 - \lambda)\Pr(C|b) + \lambda\phi_0$  from equation (9):

$$\phi_{b,E} - \phi_b = (1 - \lambda) \left[ \gamma \frac{\gamma\hat{\phi}_b}{\gamma\hat{\phi}_b + \eta(1 - \hat{\phi}_b)} + (1 - \gamma) \frac{(1 - \gamma)\hat{\phi}_b}{(1 - \gamma)\hat{\phi}_b + (1 - \eta)(1 - \hat{\phi}_b)} - \hat{\phi}_b \right],$$

which implies that

$$\phi_{b,E} - \phi_b = \frac{(1 - \lambda)\hat{\phi}_b(1 - \hat{\phi}_b)^2(\gamma - \eta)^2}{\eta(1 - \eta) + \hat{\phi}_b(\gamma - \eta)[1 - 2\eta - \hat{\phi}_b(\gamma - \eta)]} \geq 0, \quad (12)$$

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

By equations (7) and (8), it is sufficient to show the following two claims.

*Claim 1)*  $X^{NS} \geq X^S$  for all  $\phi$ .

Subtract these expressions:

$$X^{NS} - X^S = (1 - \beta)[p(\phi_{b,N}) - p(\phi_b)] - (1 - \alpha)[p(\phi_{b,E}) - p(\phi_b)],$$

which is non-negative, since  $\alpha > \beta$  by assumption and  $p(\phi_{b,N}) \geq p(\phi_{b,E})$  for all feasible  $\phi$ , from step 1 (equation 11) and monotonicity of  $p(\phi)$ .

*Claim 2)*  $Y^{NS} \geq Y^S$  for all  $\phi$ .

Subtract these expressions:

$$Y^{NS} - Y^S = \alpha[Y_g^{NS} - Y_g^S] + (1 - \alpha)[Y_b^{NS} - Y_b^S],$$

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

$$Y_k^{NS} - Y_k^S = ((1 - \beta)\rho - (1 - \alpha)\gamma)[V(\phi_{bk}^s) - V(\phi_{bk}^f)] - (\alpha - \beta)[V(\phi_{bk}) - V(\phi_{bk}^f)] \geq 0.$$

Because of the monotonicity of  $V(\phi)$  in  $\phi$  and since, by equation (4),  $\phi_b^s \geq \phi_b$ , then  $V(\phi_{bk}^s) \geq V(\phi_{bk}) \geq V(\phi_{bk}^f)$ . A sufficient condition for non-negativity is then  $((1 - \beta)\rho - (1 - \alpha)\gamma) \geq (\alpha - \beta)$ , which is the same as

$$\rho \geq \rho^* = 1 - \frac{(1 - \alpha)}{(1 - \beta)}(1 - \gamma), \quad (13)$$

where  $\frac{(1 - \alpha)}{(1 - \beta)}$  is a measure of the relative capability of experts to achieve good production results when compared to nonexperts. Hence, whenever the sufficient condition  $\rho \geq \rho^*$  holds, regardless of the value function, the likelihood of having an efficient situation reaches its maximum without scapegoating. ■

Scapegoating is a reputational buffer for competent managers. They can exploit differences in blaming abilities as an additional channel to distinguish themselves from inept managers. First, competent managers can blame workers easier than inept managers, since the workers are indeed the responsible for the failures that generate a reduction in reputation. Second, competent managers can blame nonexperts easier than experts, since nonexperts are indeed more likely to provide a bad service ( $\rho - \gamma$ ), which introduces additional gains of hiring non-experts (additional to low wages) by increasing the probability of a successful scapegoating.

To interpret the sufficient condition for the inefficiency of scapegoating,  $\rho \geq \rho^*$ , recall that  $\rho^*$  depends on  $\frac{(1 - \alpha)}{(1 - \beta)}$ , a measure of the relative likelihood experts provide a bad service when compared to nonexperts. If hiring experts almost guarantees a good service ( $\alpha \rightarrow 1$ ), then

$\rho^* \rightarrow 1$ . If hiring experts does not significantly increase the probability of a good service ( $\beta \rightarrow \alpha$ ), then  $\rho^* \rightarrow \gamma$ . Hence, the sufficient condition  $\rho \geq \rho^*$  is more difficult to hold when hiring experts is really beneficial from a productive point of view and managers can signal their competence directly in the first stage. Contrarily, if hiring experts does not make an important difference in production, competent managers tend to rely more on the use of scapegoating to signal competence, thus contributing to inefficiency.<sup>11</sup>

Finally, it is relevant to highlight at this point that I am not allowing managers taking credit for a good service exactly because I am interested in analyzing the efficiency effects of "nested" activities, which are restricted to occur only after certain outcomes in production. However, as we discussed for the case  $\gamma < \eta$ , taking credit for good services would be irrelevant for efficiency if competent managers do not have any advantage in obtaining recognition for good services. Otherwise we just have reports happening both after good and bad services, and the net effect on efficiency would be a combination between the negative effects discussed here and the positive ones we will discuss in Section 4.

### 3.3 The Intuition Behind the Inefficiency of Scapegoating

Scapegoating is valuable to competent managers to buffer reputation losses but it is irrelevant to consumers. By assumption, managers care only about reputation, and their decisions react more to activities that better maintain or construct reputation. If those activities occur only after particular situations, such as scapegoating occurs only after the provision of a bad service, managers try to maximize those situations likelihood, even when they are detrimental to activities that really matter to consumers.

Assume, for example, the extreme case in which  $\eta = 0$  and  $\gamma = \rho = 1$  (i.e., inept managers can never blame workers successfully, since workers are indeed not responsible, while competent managers can always blame workers successfully, since workers are indeed responsible). In this situation consumers learn whether the manager is competent or not immediately after scapegoating. Here, blaming is more informative than production about the manager's type, then competent managers prefer to provide a bad service, which is an excuse to signal their type more effectively through scapegoating than through production. In this extreme example, competent managers never hire experts, since nonexperts are costless and increase the probability of scapegoating.

Naturally, the previous extreme example is consistent with the case  $\gamma > \eta$ . But what happens if  $\gamma \leq \eta$ ? Even when this assumption is not natural in our setting (because it implies inept

---

<sup>11</sup>The sufficient condition is relevant only because we do not know the shape of the value function. However, if the value function is linear (risk neutrality) or convex on  $\phi$ , scapegoating would always imply inefficiency, regardless of the specific value of  $\rho$ .

managers can blame workers easier than competent managers, even when inept managers are the real responsible for bad services and competent managers are not), it is informative about the mechanism at play. As shown formally, in this case blaming does not have any effect on efficiency conditions. This result is a version of a cheap talk game. Inept managers can always pool probabilistically with competent managers in their blaming strategies and scapegoating cannot modify reputation further.

To gain intuition about the propositions and proofs we need to compare the reputation that competent managers expect to obtain from hiring experts and from hiring nonexperts. Without scapegoating, the reputation conditional on the first round's results is known and given by  $\phi_g$  after providing a good service and  $\phi_b$  after providing a bad service. With scapegoating, the reputation after a good service is also given by  $\phi_g$  but the expected reputation after a bad service depends on the hiring decision, since blaming nonexperts is easier. Specifically, the expected reputation after a bad service is  $\phi_{b,E}$  if hiring experts (equation (9)) and  $\phi_{b,N}$  (equation (10)) if hiring nonexperts.

The differential gains in reputation expected from providing a good service determine the incentives to hire experts. These gains are given by  $(\phi_g - \phi_b)$  without scapegoating (regardless of the hiring decision), by  $(\phi_g - \phi_{b,E})$  with scapegoating if hiring experts, and by  $(\phi_g - \phi_{b,N})$  with scapegoating if hiring nonexperts. Hence, to understand hiring incentives we need to understand how  $\phi_{b,N}$ ,  $\phi_{b,E}$ , and  $\phi_b$  relate to each other. As shown in steps 1 and 2 of the proof for Proposition 3, when  $\rho > \gamma > \eta$ , there is a clear ordering between these expressions.

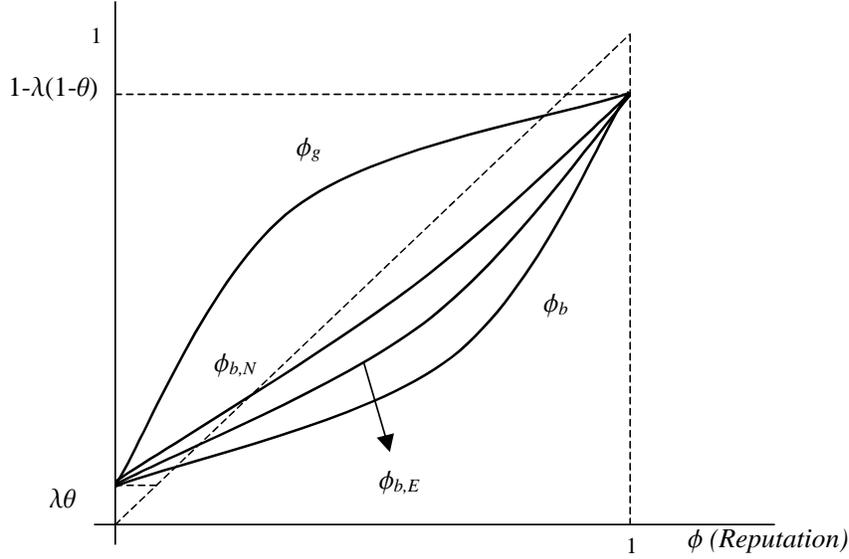
$$\phi_g > \phi_{b,N} > \phi_{b,E} > \phi_b \quad (14)$$

Graphically, Figure 1 shows the reputation posterior for each case and possible prior  $\phi$ . If the prior is  $\phi = 0$ , the update in all cases is  $\lambda\phi_0$ . If the prior is  $\phi = 1$ , the update is  $1 - \lambda(1 - \phi_0)$ . For all other values  $\phi \in (0, 1)$ , reputation updates have the ordering given by inequality (14).

The difference  $(\phi_{b,E} - \phi_b)$  in equation (11) can be interpreted as the reduction in the incentives to hire experts, and the difference  $(\phi_{b,N} - \phi_{b,E})$  in equation (12) as the increase in the incentives to hire nonexperts. While the expression  $(\gamma - \eta)$  in the numerator of equations (11) and (12) shows the magnitude of reputation maintenance due to scapegoating, the expression  $(\rho - \gamma)$  in the numerator of (11) shows the additional benefits from hiring nonexperts by taking advantage of scapegoating.

The following lemmas show that both  $(\phi_{b,N} - \phi_{b,E})$  and  $(\phi_{b,E} - \phi_b)$  are not only positive (as shown in steps 1 and 2 of the proof for Proposition 3) but also depend positively on the relative probability of scapegoating successfully.

Figure 1: Expected reputation after the first round (with and without scapegoating)



**Lemma 1** *The difference between expected reputation after a bad service from hiring experts versus hiring nonexperts ( $\phi_{b,N} - \phi_{b,E}$ ) is non decreasing in  $(\rho - \gamma)$  and in  $(\gamma - \eta)$ .*

**Proof.** Taking the derivative of expression  $(\phi_{b,N} - \phi_{b,E})$  in equation (11) with respect to  $(\rho - \gamma)$  and with respect to  $(\gamma - \eta)$ ,

$$\frac{\partial[\phi_{b,N} - \phi_{b,E}]}{\partial(\rho - \gamma)} = \frac{(1 - \lambda)\hat{\phi}_b(1 - \hat{\phi}_b)(\gamma - \eta)}{\eta(1 - \eta) + \hat{\phi}_b(\gamma - \eta)[1 - 2\eta - \hat{\phi}_b(\gamma - \eta)]^2} \geq 0$$

$$\frac{\partial[\phi_{b,N} - \phi_{b,E}]}{\partial(\gamma - \eta)} = \frac{(1 - \lambda)\hat{\phi}_b(1 - \hat{\phi}_b)(\rho - \gamma)[\eta(1 - \eta) + \hat{\phi}_b(\gamma - \eta)^2]}{[\eta(1 - \eta) + \hat{\phi}_b(\gamma - \eta)[1 - 2\eta - \hat{\phi}_b(\gamma - \eta)]^2} \geq 0.$$

The two expressions are strictly positive when  $\rho > \gamma > \eta$  and  $\hat{\phi}_b \in (0, 1)$ . ■

**Lemma 2** *The difference between expected reputation after a bad service in cases with and without scapegoating,  $(\phi_{b,E} - \phi_b)$ , is non decreasing in  $(\gamma - \eta)$ .*

**Proof.** For this proof, consider only the difference  $\phi_{b,E} - \phi_b$  in equation (12) since, as shown in Lemma 1,  $\frac{\partial[\phi_{b,N} - \phi_{b,E}]}{\partial(\gamma - \eta)} \geq 0$ . Taking derivatives of  $(\phi_{b,E} - \phi_b)$  with respect to  $(\gamma - \eta)$ ,

$$\frac{\partial[\phi_{b,E} - \phi_b]}{\partial(\gamma - \eta)} = \frac{(1 - \lambda)\hat{\phi}_b(1 - \hat{\phi}_b)^2(\gamma - \eta)[2\eta(1 - \eta) + (1 - 2\eta)\hat{\phi}_b(\gamma - \eta)]}{[\eta[(1 - \eta) + \hat{\phi}_b(\gamma - \eta)[1 - 2\eta - \hat{\phi}_b(\gamma - \eta)]]^2} \geq 0,$$

which is non-negative because in the numerator,  $(1 - \eta) \geq (\gamma - \eta) \geq \hat{\phi}_b(\gamma - \eta)$ . This is also strictly positive whenever  $\rho > \gamma > \eta$  and  $\hat{\phi}_b \in (0, 1)$ . ■

The difference in the blaming abilities between competent managers and inept managers ( $\gamma - \eta$ ) basically measures the drop in expected reputation that, because of scapegoating, does not occur after a bad service. Hence, an increase in  $(\gamma - \eta)$  not only reduces the incentives to hire experts (by increasing  $\phi_{b,E} - \phi_b$ ) but also increases the incentives to hire nonexperts (by increasing  $\phi_{b,N} - \phi_{b,E}$ ).

Similarly, the difference in the abilities between blaming experts and nonexperts ( $\rho - \gamma$ ) measures the greater probability of having a successful scapegoating from hiring nonexperts. Hence, an increase in  $(\rho - \gamma)$  makes hiring nonexperts even more beneficial (by further increasing  $\phi_{b,N} - \phi_{b,E}$ ).

### 3.4 Machiavellian Effect

The next proposition shows that, with scapegoating, efficient outcomes are more likely during good times, measured as times with a higher  $\alpha$  and  $\beta$  (maintaining the relative difference  $(\alpha - \beta)$  constant), since it is more likely to provide good services in general during those times.

#### Proposition 4 Machiavellian Effect

Suppose  $\lambda \in (0, 1)$ ,  $\phi \in [\lambda\phi_0, 1 - \lambda(1 - \phi_0)]$ ,  $\delta \in (0, 1)$ ,  $\phi_0 \in (0, 1)$ ,  $\alpha' > \alpha$ ,  $(\alpha' - \beta') = (\alpha - \beta)$ , and competent managers have better blaming capabilities than inept managers ( $\gamma > \eta$ ).

When scapegoating is not allowed, efficiency depends only on  $(\alpha - \beta)$ , but not on the level of  $\alpha$  or  $\beta$ .

When scapegoating is allowed, in any equilibrium where competent managers decide to hire experts when  $\alpha$ , they also decide to hire experts when  $\alpha' > \alpha$ , while the contrary is not true. Hence, efficiency depends both on  $(\alpha - \beta)$  and the level of  $\alpha$  and  $\beta$ .

**Proof.** First I show that  $\Delta^S(\alpha') \geq \Delta^S(\alpha)$ . For this I prove that  $X^S(\alpha') + \delta(1 - \lambda)Y^S(\alpha') \geq X^S(\alpha) + \delta(1 - \lambda)Y^S(\alpha)$  for all  $\phi \in (0, 1)$ . Considering equation (8), it suffices to show the following two claims:<sup>12</sup>

<sup>12</sup>In what follows, and just to make the point cleaner, I assume consumers do not observe  $\alpha$  when updating reputation. Hence the reputation update is just based on  $E(Pr(g|I)) = E(\beta) \equiv \bar{\beta}$  and  $E(Pr(g|I)) = E(\alpha|Eff)Pr(Eff) + E(\beta|Ineff)Pr(Ineff) \equiv \bar{\alpha} > \bar{\beta}$  (this is the expected  $\alpha$  for all  $\alpha$  that sustains an efficient equilibrium and the expected  $\beta$  for all  $\alpha$  that does not sustain an efficient equilibrium).

*Claim 1)  $X^S(\alpha') \geq X^S(\alpha)$  for all  $\phi$ .*

Subtracting these expressions,

$$X^S(\alpha') - X^S(\alpha) = (\alpha' - \alpha)[p(\phi_{b,N}) - p(\phi_{b,E})],$$

which is non negative, since  $\alpha' > \alpha$  by assumption;  $p(\phi)$  is monotonic in  $\phi$  and by equation (11),  $\phi_{b,N} \geq \phi_{b,E}$ .

*Claim 2)  $Y^S(\alpha') \geq Y^S(\alpha)$  for all  $\phi$ .*

Subtract these expressions:

$$Y^S(\alpha') - Y^S(\alpha) = E(\alpha)[Y_g^S(\alpha') - Y_g^S(\alpha)] + (1 - E(\alpha))[Y_b^S(\alpha') - Y_b^S(\alpha)],$$

is non negative, since  $Y_k^S(\alpha') - Y_k^S(\alpha) = (\alpha' - \alpha)(\rho - \gamma)[V(\phi_{bk}^s) - V(\phi_{bk}^f)] \geq 0$  for  $k \in \{g, b\}$ . This is because  $\alpha' > \alpha$ ,  $\rho \geq \gamma$  by assumption,  $V(\phi)$  is monotonic in  $\phi$  and  $\phi_b^s \geq \phi_b^f$ . Assuming scapegoating in the future does not change the conclusion. Hence, in times with a higher  $\alpha$ , competent managers hire experts for a wider range of wages  $w$ , and efficiency is more likely.

Finally I show that, when there is no scapegoating, efficiency conditions only depend on  $(\alpha - \beta)$  and not on the level of  $\alpha$  or  $\beta$ , as long as the difference  $(\alpha - \beta)$  remains constant. From Proposition 1, recall  $X^{NS}(\alpha') - X^{NS}(\alpha) = 0$  and  $Y^{NS}(\alpha') - Y^{NS}(\alpha) = 0$ . Then,  $\Delta^{NS}(\alpha') = \Delta^{NS}(\alpha)$ , as long as  $(\alpha' - \beta') = (\alpha - \beta)$ . ■

This Machiavellian effect arises because, when it is more difficult to provide a good quality service, managers are more concerned about potential reputation losses. Without scapegoating, this concern is irrelevant: even when the probability of providing a bad service is big, the differences in probabilities  $(\alpha - \beta)$  from hiring experts or nonexperts is what determines hiring decisions. With scapegoating, however, hiring nonexperts become more attractive because they are easier to blame, buffering the reputation loss if a bad service in fact is provided, exactly as proposed by Machiavelli.

## 4 Cheap Ways to Achieve an Efficient Equilibrium

We discussed how scapegoating generates inefficiencies by being used after failures (in our case, the provision of a bad service). A natural question arises. What happens with activities that are used exclusively after successes? It is possible to find many examples of this kind of activities. In the sporting arena, All-Star Games, national teams and international championships (such as the Soccer World Cup) are organized for the participation of the best

players. In organizations, corporations, and public offices, additional funds and responsibilities are assigned to divisions that outperform. In academic environments, round tables and plenary sessions at professional meetings are held by top researchers. In the entertainment industry, TV shows invite successful music and movie stars to exhibit their charisma or talent.

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

#### 4.1 Efficiency of Nested Activities after Successes

The model can be reinterpreted and modified to introduce activities after the provision of a good service. Assume that instead of an irrelevant activity nested after a bad service, such as scapegoating, the game is characterized by an irrelevant activity nested after a good service, such as providing a complementary product or following up with the client.

I assume the structure of probabilities, timing, and parameters for the production stage have the same interpretation as before. The difference appears in the second stage. After a bad service the game ends, but after a good service there is a nested activity, which can succeed  $(g, s)$  or fail  $(g, f)$  (these are basically in the same spirit as  $(b, s)$  and  $(b, f)$  in the scapegoating setting). In particular, there may be differences in the probabilities of being successful at the nested activity,  $\Pr(s|I, g, N) = \eta_g$ ,  $\Pr(s|C, g, E) = \gamma_g$  and  $\Pr(s|C, g, N) = \rho_g$ .<sup>13</sup>

For example, if  $\gamma_g > \rho_g$ , hiring experts increases the probability of being successful not only at producing but also at the additional nested activity. Competent managers can choose the probability of exerting efforts in performing the nested activity. In the same vein as scapegoating, if the manager exerts efforts in the nested activity, the probability of success is  $\eta_g$ ,  $\gamma_g$  or  $\rho_g$ , depending on the type of manager and worker. If the manager does not exert efforts, the probability of success is 0, but the consumers do not know whether the failure happens with exerting efforts or not. In what follows, I assume  $\eta_g \leq \rho_g \leq \gamma_g \leq 1$ . Note this is not the same ordering we assumed for the scapegoating case. In this case it makes more sense to assume that experts increase the probability of success in the nested activity, contrary to the case of scapegoating in which experts reduce the probability of success in blaming. In

---

<sup>13</sup>Many of the examples discussed do not need delegation. In fact, in a better description competent managers would decide between exerting high or low efforts, while inept managers would only be able to exert low efforts. This alternative environment, although the same in spirit, is different in that managers would need to choose the effort level both before the first and the second rounds and pay twice the effort costs. Introducing this modification does not change the main conclusion, though.

any case, using the ordering  $\eta_g \leq \gamma_g \leq \rho_g \leq 1$ , do not modify the efficiency results since the important assumption is  $\eta_g \leq \gamma_g$ .

The equilibrium definition in this environment is the same as before, except that the three possible belief updates (replacing equations (3), (4), and (5)) are  $\phi_b$  (after a bad service),  $\phi_g^s$  (after a good service and a success), and  $\phi_g^f$  (after a good service and a failure). It is straightforward to show that in the efficient equilibrium, for all  $\phi$ , after the first round  $\phi_g > \phi > \phi_b$  and after the potential second round  $\phi_g^s > (<)\phi_g > (<)\phi_g^f$  if  $\gamma_g > (<)\eta_g$ . As in the scapegoating situation, only when  $\gamma_g > \eta_g$  may the nested stage generate a new reputation updating and affect efficiency.<sup>14</sup>

As in equations (11) and (12), it is also possible to define expected reputation after good results in the case of hiring experts and in the case of hiring nonexperts. Equations (11) and (12) could be restated as

$$\phi_{g,E} - \phi_{g,N} = \frac{(1 - \lambda)\hat{\phi}_g(1 - \hat{\phi}_g)(\gamma_g - \rho_g)(\gamma_g - \eta_g)}{\eta_g(1 - \eta_g) + \hat{\phi}_g(\gamma_g - \eta_g)[1 - 2\eta_g - \hat{\phi}_g(\gamma_g - \eta_g)]} > 0 \quad (15)$$

$$\phi_{g,E} - \phi_g = \frac{(1 - \lambda)\hat{\phi}_g(1 - \hat{\phi}_g)^2(\gamma_g - \eta_g)^2}{\eta_g(1 - \eta_g) + \hat{\phi}_g(\gamma_g - \eta_g)[1 - 2\eta_g - \hat{\phi}_g(\gamma_g - \eta_g)]} > 0, \quad (16)$$

where  $\hat{\phi}_g = \Pr(C|g)$ .

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

$$\phi_{g,N} - \phi_g = \frac{(1 - \lambda)(\hat{\phi}_g)(1 - \hat{\phi}_g)(\gamma_g - \eta_g)[(\rho_g - \eta_g) - \hat{\phi}_g(\gamma_g - \eta_g)]}{\eta_g(1 - \bar{x}_g) + \hat{\phi}_g(\gamma_g - \eta_g)[1 - 2\eta_g - \hat{\phi}_g(\gamma_g - \eta_g)]}, \quad (17)$$

A clear ordering exists among these expressions when  $\gamma_g > \eta_g$ , as shown in Figure 2.

$$\phi_{g,E} > \phi_{g,N} > \phi_b \quad \text{and} \quad \phi_{g,E} > \phi_g > \phi_b \quad (18)$$

and, even when not relevant for our results,

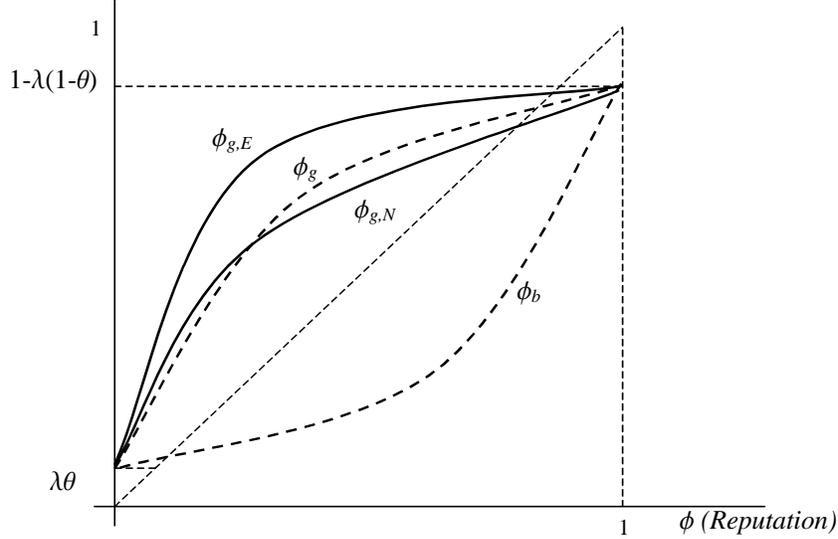
$$\phi_{g,N} > \phi_g \quad \text{when} \quad \hat{\phi}_g < \frac{\rho_g - \eta_g}{\gamma_g - \eta_g}.$$

Hiring experts increases expected reputation after a good service (equation (16) is positive) while hiring nonexperts decreases expected reputation after a good service (equation (17) is

<sup>14</sup>If  $\gamma_g \leq \eta_g$ , inept managers prefer to pool with competent managers in order to be confused with them. In this way, inept managers would not signal their own ineptitude (this is the same logic explained in Section 3.3).

negative) for relatively high reputation levels. The net effect is always an increase in expected reputation gains from hiring experts (equation (15) is positive).

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



While Figure 2 delivers the basic intuition that sustains efficiency from nested activities after a good service when  $\gamma_g > \eta_g$ , formal proofs are very similar to the proofs for Propositions 1 and 2. Contrary to scapegoating, here it makes more sense to assume  $\gamma_g > \rho_g$ . If this is the case, the irrelevant stage after successes always increases the likelihood of achieving an efficient equilibrium, without requiring a sufficient condition.

Summarizing, activities nested after a good service align and exploit reputation concerns to achieve efficient outcomes, even when these activities are completely irrelevant and costless.<sup>15</sup> The main difference between nested activities after failures (such as scapegoating) and nested activities after successes is that, while the former reduces the expected reputation losses from a bad service (from  $(\phi_g - \phi_b)$  to  $(\phi_g - \phi_{b,E})$ ), reducing the incentives to hire experts, the latter increases the expected reputation gains from a good service (from  $(\phi_g - \phi_b)$  to  $(\phi_{g,E} - \phi_b)$ ), increasing the incentives to hire experts. Furthermore, while the former increases the incentives to hire nonexperts (by  $(\phi_{b,N} - \phi_{b,E})$ ), the latter reduces them (by  $(\phi_{g,E} - \phi_{g,N})$ ).

Finally, the Machiavellian effect persists also in this case. There are more incentives to hire experts when the probability of providing a good service increases, since it increases the probability of reaching the nested activity, further increasing the incentives to hire experts.

<sup>15</sup>An extreme but illustrative example is the following. Suppose that hiring experts is efficient, but incentives from production are not enough for managers to do it. Assume also that experts heavily outperform nonexperts at playing chess. A cheap way to achieve efficiency would be to introduce a chess game right after a good service is provided!

## 5 Conclusions

Firms can exploit irrelevant activities that have good signalling properties by nesting them after a success in a non very informative production activity. This nesting pattern works because it better aligns incentives generated by agents' reputation concerns. Even though these irrelevant activities do not affect production directly, they do so indirectly through the decisions governing the probability of success in production. Similarly, allowing these irrelevant activities to happen after failures reduces incentives for agents to behave efficiently.

In this paper I discuss a model of scapegoating to illustrate these effects in a contained environment, but the same ideas are applicable in other settings. In this case, scapegoating corrupts the incentives that reputation concerns introduce, by attenuating potential losses of reputation, hence reducing the chances of costly decisions conducive to obtaining good results, such as hiring experts. Furthermore, scapegoating in fact increases the incentives to hire nonexperts in order to blame them more easily if something goes wrong. In this situation it is not only better to ban scapegoating, but it is in fact optimal to force reports only after a success in production.

There are many potential applications this model can cover. Recalls of bad products can be considered an activity nested after failures, which give firms the chance to show their competence by handling the recall well. This seemingly beneficial activity may in fact reduce the efforts put in production in the first place. Nested activities have also applications in political economy. How an executive's leader chose the quality of its ministers? How an executive's leader can induce ministers to perform without using efficiency wages, just relying on firing decisions?<sup>16</sup>

In this paper the relative wage of experts is exogenous  $w$ , however endogeneize it strengthens the results. If experts charge relatively more when it is more likely they are blamed (let's say because they have reputation concerned on their own), there are even less incentives to hire experts when scapegoating is known to be a possibility. Even when beyond the scope of this paper, it would be interesting to expand the model to study the interaction between characteristics of the labor market and organizational hierarchies and its effects on efficiency when scapegoating exists.

Finally, it is important to reiterate that it is also beyond the scope of this paper to provide a comprehensive model of scapegoating. I just highlight a novel direct negative effect of scapegoating arising from its property of being an activity typically used after failures. This is why conclusions are biased toward the assignment of complete responsibility to managers. However, there may be potential arguments in favor of relaxing accountability, such as specializa-

---

<sup>16</sup>Interesting discussions about these issues are Dowding and Dumont (2008) and Dewan and Myatt (2010).

tion and scale. Hence, a more complete model of scapegoating, considering all determinants, would be necessary to obtain the optimal allocation of responsibility and accountability to managers.

## References

- Alesina, Alberto, and Guido Tabellini. 2005. "Why Do Politicians Delegate?" Working Paper 11531. National Bureau of Economic Research.
- . 2007. "Bureaucrats or Politicians? Part I: A Single Policy Task." *American Economic Review* 97:169–179.
- . 2008. "Bureaucrats or Politicians? Part II: Multiple Policy Tasks." *Journal of Public Economics* 92:425–447.
- Alonso-Pauli, Eduard, and David Perez-Castrillo. 2010. "Codes of Best Practice in Competitive Markets for Managers." *Economic Theory*. DOI: 10.1007/s00199-010-0537-y.
- Bell, Nancy E., and Philip E. Tetlock. 1989. "The Intuitive Politician and the Assignment of Blame in Organizations." In *Impression Management in the Organization*, edited by Robert Giacalone and Paul Rosenfel, 105–123. Hillsdale, NJ: Erlbaum.
- Cappelen, Alexander, and Bertil Tungodden. 2009. "Rewarding Effort." *Economic Theory* 39:425–441.
- Cole, Hal, and Patrick Kehoe. 1996. "Reputation Spillover Across Relationships: Reviving Reputation Models of Debt." Working Paper 5486, National Bureau of Economic Research.
- Cripps, Martin, George Mailath, and Larry Samuelson. 2004. "Imperfect Monitoring and Impermanent Reputations." *Econometrica* 72:407–432.
- . 2007. "Disappearing Private Reputations in Long-Run Relationships." *Journal of Economic Theory* 134:287–316.
- Dewan, Torun, and David Myatt. 2010. "The Declining Talent Pool of Government." *American Journal of Political Science* 54 (2): 267–286.
- Dezso, Cristian. 2004. "Scapegoating and Firm Reputation." Mimeo, Stern School of Business, New York University.
- Douglas, Tom. 1995. *Scapegoating: Transferring Blame*. London: Routledge.
- Dowding, Keith, and Patrick Dumont. 2008. *The Selection of Ministers in Europe*. Routledge Advances in European Politics. Routledge, London.

- Ely, Jeffrey, and Juuso Valimaki. 2003. "Bad Reputation." *Quarterly Journal of Economics*, no. 118:785–814.
- Fudenberg, Drew, and David K. Levine. 1989. "Reputation and Equilibrium Selection in Games with a Patient Player." *Econometrica* 57:759–778.
- . 1992. "Maintaining a Reputation When Strategies Are Imperfectly Observed." *Review of Economic Studies* 59:561–579.
- Huson, Mark R., Paul H. Malatesta, and Robert Parrino. 2004. "Managerial Succession and Firm Performance." *Journal of Financial Economics* 74:237–275.
- Koch, Alexander, and Eloic Peyrache. 2008. "Mixed up? That's Good for Motivation." *Economic Theory* 34:107–125.
- Kreps, David, and Robert Wilson. 1982. "Reputation and Imperfect Information." *Journal of Economic Theory* 27:253–279.
- Levy, Gilat. 2005. "Careerist Judges and the Appeals Process." *RAND Journal of Economics*, no. 36:275–297.
- Mailath, George, and Larry Samuelson. 1998. "Your Reputation Is Who You're Not, Not Who You Would Like to Be." CARESS, Working Paper 98-11, University of Pennsylvania.
- . 2001. "Who Wants a Good Reputation?" *Review of Economic Studies* 68 (2): 415–441.
- . 2006. *Repeated Games and Reputations*. New York: Oxford University Press.
- Martin, John. 1997. "Changing Accountability Relations: Politics, Consumers and the Market." Public Management Service, Organisation for Economic Cooperation and Development.
- Miettinen, Topi. 2010. "Paying Attention to Payoffs in Analogy-Based Learning." *Economic Theory*. DOI: 10.1007/s00199-010-0565-7.
- Milgrom, Paul, and John Roberts. 1982. "Predation, Reputation and Entry Deterrence." *Journal of Economic Theory* 27:280–312.
- Mylovanov, Tymofiy, and Patrick Schmitz. 2008. "Task Scheduling and Moral Hazard." *Economic Theory* 37:307–320.
- Polidano, Charles. 1999. "The Bureaucrat Who Fell Under a Bus: Ministerial Responsibility and the Derek Lewis Affair in Britain." *Governance: An International Journal of Policy, Administration and Institutions*. 12:201–229.
- Segendorff, Bjorn. 2000. "A Signalling Theory of Scapegoats." Mimeo, Department of Economics, Stockholm School of Economics.
- Tadelis, Steven. 2003. "Firm Reputation with Hidden Information." *Economic Theory* 21:635–651.

Winter, Eyal. 2001. "Scapegoats and Optimal Allocation of Responsibilities." Mimeo, Center for Rationality and Interactive Decision Theory, The Hebrew University of Jerusalem.

## A Appendix

**Proof for Proposition 1.** Fix  $\phi$  and suppose an efficient situation (i.e., competent managers always choose to hire experts,  $\tau(\phi) = 1$ ). It is straightforward to show that for all  $\phi$ , after the first round  $\Pr(C|g) > \Pr(C) > \Pr(C|b)$ .

Hence, for all feasible  $\phi$ ,  $\varphi(\varphi(\phi|g)|g) = \phi_{gg} > \phi_g > \phi > \phi_b > \phi_{bb}$  and  $\phi_{gk} > \phi_{bk}$  for  $k \in \{g, b\}$ .

Competent managers' value function when hiring experts is

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

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

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

Thus,

$$V(\phi, E) - V(\phi; N) = -w + \delta(1 - \lambda)X^{NS} + \delta^2(1 - \lambda)^2Y^{NS},$$

where

$$\begin{aligned} X^{NS} &= (\alpha - \beta)[p(\phi_g) - p(\phi_b)] \\ Y^{NS} &= (\alpha - \beta)[\alpha(V(\phi_{gg}) - V(\phi_{bg})) + (1 - \alpha)(V(\phi_{gb}) - V(\phi_{bb}))] \end{aligned}$$

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

$$w \leq \delta(1 - \lambda)[X^{NS} + \delta(1 - \lambda)Y^{NS}]; \quad \text{for all } \phi \in [\lambda\phi_0, 1 - \lambda(1 - \phi_0)],$$

Then we can define  $\Delta^{NS}$  as the minimum value of the expression  $\delta(1 - \lambda)[X^{NS} + \delta(1 - \lambda)Y^{NS}]$  over the range  $\phi \in [\lambda\phi_0, 1 - \lambda(1 - \phi_0)]$

$$\Delta^{NS} = \min_{\phi \in [\lambda\phi_0, 1 - \lambda(1 - \phi_0)]} \{ \delta(1 - \lambda)[X^{NS} + \delta(1 - \lambda)Y^{NS}] \}. \quad (19)$$

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

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

- $X^{NS} = (\alpha - \beta)[p(\phi_g) - p(\phi_b)] > 0$ , since  $\alpha > \beta$  and  $p(\phi)$  is monotonically increasing in  $\phi$  ( $\frac{\partial p(\phi)}{\partial \phi} = \alpha - \beta > 0$ ).<sup>17</sup>
- $Y^{NS} > 0$ , since  $\alpha > \beta$ ,  $V(\phi_{gg}) > V(\phi_{bg})$  and  $V(\phi_{gb}) > V(\phi_{bb})$  (the value function  $V$  is monotonically increasing in  $\phi$ ).<sup>18</sup>

■

**Proof for Proposition 2.** This proof proceeds in two steps. First, I solve for managers' scapegoating decisions,  $\nu_j^i(\phi)$ , consistent with consumers' beliefs. Second, using these results from the scapegoating stage, we derive conditions for an efficient equilibrium.

### Step 1: Scapegoating stage

First, take as given consumers' beliefs about scapegoating probabilities and determine optimal blaming decisions by inept managers working with nonexperts and competent managers working with experts ( $\nu_N^I(\phi)$  and  $\nu_E^C(\phi)$ ). These are the two relevant probabilities for consumers to update beliefs, since we are focusing only on efficient equilibria in which competent managers always hire experts. Second, considering the optimal scapegoating, we check if beliefs are correct and consistent with those strategies.

When competent managers scapegoat with probability  $\nu_E^C(\phi)$  in the efficient equilibrium,

$$V^C(\phi, \nu_E^C) = p(\phi) - w + \delta(1 - \lambda) \left[ \alpha V^C(\phi_g) + (1 - \alpha) \left[ \nu_E^C \gamma V^C(\phi_b^s) + (1 - \nu_E^C \gamma) V^C(\phi_b^f) \right] \right],$$

where  $\phi_b^s$  and  $\phi_b^f$  are given by consumers' beliefs  $\nu_N^I(\phi)$  and  $\nu_E^C(\phi)$ .

For competent managers, any deviation from  $\nu_E^C$ , say to  $\nu_E^{C'}$ , we can define

$$VD^C(\nu_E^C) = V^C(\phi, \nu_E^{C'}) - V^C(\phi; \nu_E^C) = \delta(1 - \lambda)(1 - \alpha)(\nu_E^{C'} - \nu_E^C)\gamma[V^C(\phi_b^s) - V^C(\phi_b^f)].$$

Similarly, we can construct this differential for inept managers that deviate from  $\nu_E^I$  to  $\nu_E^{I'}$ ,

$$VD^I(\nu_E^I) = V^I(\phi, \nu_E^{I'}) - V^I(\phi; \nu_E^I) = \delta(1 - \lambda)(1 - \beta)(\nu_E^{I'} - \nu_E^I)\eta[V^I(\phi_b^s) - V^I(\phi_b^f)].$$

**1)** Assume consumers believe  $\nu_E^C \gamma = \nu_N^I \eta$ . By equations (4) and (5),  $\phi_b^s = \phi_b^f = \phi_b$ . Since  $V(\phi_b^s) - V(\phi_b^f) = 0$ , competent managers are indifferent between choosing any  $\nu_E^C \in [0, 1]$  (regardless of  $(\nu_E^{C'} - \nu_E^C)$ , always  $VD(\nu_E^C) = 0$ ). Similarly, inept managers are indifferent between choosing any  $\nu_N^I \in [0, 1]$  (regardless of  $(\nu_N^{I'} - \nu_N^I)$ ,  $VD(\nu_N^I) = 0$ ). Hence, consumers'

<sup>17</sup>More specifically, as in Mailath and Samuelson (2001), suppose  $F$  and  $G$  are two distributions describing consumers beliefs over the delegation decisions by competent managers in period  $t$ . If  $F$  first-order stochastically dominates  $G$ , then managers' revenues in period  $t$  under  $F$  are higher than they are under  $G$ .

<sup>18</sup>Following Mailath and Samuelson (2001), let  $f_t(\phi, \phi_0, t_0)$  be the distribution of consumers' posteriors  $\phi$  at time  $t > t_0$  induced by strategy  $\tau$  given period- $t_0$  posteriors  $\phi_0$ . 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.

beliefs  $\nu_E^C \gamma = \nu_N^I \eta$  are correct and consistent with equilibrium strategies, supporting multiple pooling equilibria in which no further reputation update is obtained after scapegoating.

2) Assume consumers believe  $\nu_E^C \gamma > \nu_N^I \eta$ . Then,  $\phi_b^s > \phi_b > \phi_b^f$ . Since  $V(\phi_b^s) - V(\phi_b^f) > 0$ , competent managers choose  $\nu_E^{C'} = 1$ , which maximizes  $VD(\nu_E^C)$ . Similarly, inept managers will choose  $\nu_N^{I'} = 1$ . Only consumers' beliefs  $\nu_E^C = 1$  and  $\nu_N^I = 1$  will be correct, which are consistent with beliefs  $\nu_E^C \gamma > \nu_N^I \eta$  only when  $\gamma > \eta$ . This is the only separating equilibrium in which scapegoating represents an additional reputation updating. If  $\gamma < \eta$  this is not an equilibrium.

3) Assume consumers believe  $\nu_E^C \gamma < \nu_N^I \eta$ . Then,  $\phi_b^s < \phi_b < \phi_b^f$ . Since  $V(\phi_b^s) - V(\phi_b^f) < 0$ , competent managers choose  $\nu_E^{C'} = 0$  and inept managers  $\nu_N^{I'} = 0$ . Only consumers' beliefs  $\nu_E^C = 0$  and  $\nu_N^I = 0$  will be correct, which is not consistent with beliefs in which  $\nu_E^C \gamma < \nu_N^I \eta$ . This case cannot be an equilibrium.<sup>19</sup>

Since scapegoating is free, generally it cannot be used to improve reputation unless  $\nu_N^I = 1$ ,  $\nu_E^C = 1$  and  $\gamma > \eta$ , which is the only separating equilibrium.

## Step 2: Hiring stage

1) Let  $\gamma \leq \eta$ .

Fix  $\phi$  and suppose an efficient situation,  $\tau(\phi) = 1$ . Since the only possible equilibrium under the scapegoating stage is a pooling one, where  $\phi_b^s = \phi_b^f = \phi_b$ , we have exactly the same expressions used to obtain equilibrium conditions without scapegoating (in the proof for Proposition 1). Hence, under  $\gamma \leq \eta$ , scapegoating does not affect efficiency conditions.

2) Let  $\gamma > \eta$ .

Even when pooling equilibria in scapegoating that do not affect efficiency conditions exist in this situation, we will focus on the unique separating equilibrium in which  $\nu_N^I = 1$ ,  $\nu_E^C = 1$  and  $\nu_N^C = 1$  such that  $\phi_b^s > \phi_b > \phi_b^f$ .

Fix  $\phi$  and suppose an efficient situation,  $\tau(\phi) = 1$ . It is straightforward to show that for all  $\phi$ , after the first round  $\Pr(C|g) > \Pr(C) > \Pr(C|b)$  and after the potential second round  $\Pr(C|b, s) > \Pr(C|g) > \Pr(C|b, f)$  if  $\nu_E^C \gamma > \nu_N^I \eta$ . For all feasible  $\phi$ ,  $\varphi(\varphi(\phi|g)|g) = \phi_{gg} > \phi_g > \phi > \phi_b > \phi_{bb}$  and  $\phi_{gk} > \phi_{bsk} > \phi_{bk} > \phi_{bfk}$  for  $i \in \{g, b_s, b_f\}$ .

Competent managers' value function when hiring experts is

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

and when deviating, and hiring nonexperts is

$$V(\phi, N) = p(\phi) + \delta(1 - \lambda)[\beta V(\phi_g) + (1 - \beta)[\rho V(\phi_b^s) + (1 - \rho)V(\phi_b^f)]].$$

<sup>19</sup>Because we are focusing on efficient equilibria, we checked beliefs for  $\nu_E^C$  and  $\nu_N^I$ , but a competent type that deviated in the first stage hiring nonexperts will also choose any  $\nu_N^C \in [0, 1]$  in 1),  $\nu_N^C = 1$  in 2), and  $\nu_N^C = 0$  in 3).

Thus,

$$V(\phi, E) - V(\phi; N) = -w + \delta(1 - \lambda)X^S + \delta^2(1 - \lambda)^2Y^S$$

where

$$X^S = (\alpha - \beta)p(\phi_g) + (1 - \alpha)[\gamma p(\phi_b^s) + (1 - \gamma)p(\phi_b^f)] - (1 - \beta)[\rho p(\phi_b^s) + (1 - \rho)p(\phi_b^f)]$$

$$Y^S = \alpha Y_g^S + (1 - \alpha)[\gamma Y_{b_s}^S + (1 - \gamma)Y_{b_f}^S]$$

and, for  $k \in \{g, b_s, b_f\}$ ,

$$Y_k^S = (\alpha - \beta)V(\phi_{gk}) + (1 - \alpha)[\gamma V(\phi_{bk}^s) + (1 - \gamma)V(\phi_{bk}^f)] - (1 - \beta)[\rho V(\phi_{bk}^s) + (1 - \rho)V(\phi_{bk}^f)]$$

It is useful to express  $X^S$  in terms of expected reputation after the hiring decision defining

$$\phi_{b,E} = \gamma\phi_b^s + (1 - \gamma)\phi_b^f \quad (20)$$

$$\phi_{b,N} = \rho\phi_b^s + (1 - \rho)\phi_b^f. \quad (21)$$

Hence, exploiting the linearity of  $p(\phi)$ ,<sup>20</sup>

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

An equilibrium in which competent managers only hire experts requires that  $V(\phi, E) - V(\phi; N) \geq 0$  for all feasible reputations  $\phi$ . A necessary condition is,

$$w \leq \delta(1 - \lambda)[X^S + \delta(1 - \lambda)Y^S]; \quad \text{for all } \phi \in [\lambda\phi_0, 1 - \lambda(1 - \phi_0)],$$

Then we can define  $\Delta^S$  as the minimum value of the expression  $\delta(1 - \lambda)[X^S + \delta(1 - \lambda)Y^S]$  over the range  $\phi \in [\lambda\phi_0, 1 - \lambda(1 - \phi_0)]$

$$\Delta^S = \min_{\phi \in [\lambda\phi_0, 1 - \lambda(1 - \phi_0)]} \{ \delta(1 - \lambda)[X^S + \delta(1 - \lambda)Y^S] \} \quad (22)$$

To save notation, it is possible to assume a case in which the future does not have scapegoating possibilities, so there is just one current shot at blaming. In this case, from tomorrow on, it would be possible to have only two possible states  $k \in \{g, b\}$ . It is straightforward to check that (22) is simplified to  $Y^S = \alpha Y_g^S + (1 - \alpha)Y_b^S$ . This last expression is the one used in Proposition 2. ■

**Extension of proof for Proposition 3.** Assuming scapegoating in current and future periods, only modifies the definition of  $Y^S$ , in equation (22).

<sup>20</sup>It is not possible to do the same for  $Y_k^S$  because we do not know the form of the value functions, just their monotonicity in  $\phi$  (recall we are not imposing linearity of  $V(\phi)$ ).

Hence, we need to prove  $Y^{NS} - Y^S \geq 0$ . Then,

$$Y^{NS} - Y^S = \alpha(Y_g^{NS} - Y_g^S) + (1 - \alpha)(Y_b^{NS} - \gamma Y_{b_s}^S - (1 - \gamma)Y_{b_f}^S)$$

This expression will be non-negative whenever  $Y_g^{NS} - Y_g^S \geq 0$  (proved in Proposition 3) and  $Y_b^{NS} - \gamma Y_{b_s}^S - (1 - \gamma)Y_{b_f}^S \geq 0$ , which happens under two sufficient conditions,

**1)** Since  $\phi_{gb_s} \geq \phi_{gb} \geq \phi_{gb_f}$  and  $(\alpha - \beta) [\gamma V(\phi_{gb}) + (1 - \gamma)V(\phi_{gb})] \geq (\alpha - \beta) [\gamma V(\phi_{gb_s}) + (1 - \gamma)V(\phi_{gb_f})]$ ,

$$\gamma \leq \frac{V(\phi_{gb}) - V(\phi_{gb_f})}{V(\phi_{gb_s}) - V(\phi_{gb_f})} \leq 1,$$

**2)**  $[(1 - \beta)\rho - (1 - \alpha)\gamma][V(\phi_{bb_s}^s) - V(\phi_{bb_s}^f)] \geq (\alpha - \beta)[V(\phi_{bb}) - V(\phi_{bb_s}^f)]$ .

Given  $V(\phi_{bb_s}^s) > V(\phi_{bb_s}^f)$ , the sufficient condition for this to hold is (as in (13))

$$\rho > \rho^* = 1 - \frac{(1 - \alpha)}{(1 - \beta)}(1 - \gamma).$$

■