PRELIMINARY WORK

Incentives, procedure and the extra mile

Antoine Prévet*

Abstract

We model the widespread intuition that more bureaucratic management can lead to less effort and quality in a principal-agent framework. Introducing a procedure aiming at codifying and easing verification of effort can be socially inefficient yet chosen by the principal. In the model with three performance levels, the agent can either shirk, work or go the extra mile, standard effort decreases the occurrence of bad performance, while the extra mile increases the one of the best and is harder to codify. The usual monotone likelihood ratio property naturally breaks down in such a setting. We show that the procedure makes the extra mile both cheaper and less implemented than without the procedure. The introduction of the procedure has implications for organizational design where the principal faces a trade off between incentive gains from task bundling and verification.

JEL classification: D23, D82.

Keywords: Moral Hazard, Limited Liability, Procedure, Monitoring.

*Université Paris 1 Panthéon Sorbonne – Paris School of Economics. E-mail: antoine.prevet@univ-paris1.fr. Postal Address: 106-112 Boulevard de l'Hôpital, 75013 PARIS France.

Contents

1 Introduction

In most activities, the way things are done matters as much as the final outcome to organization heads. In addition, the actions of employees are often not observable, giving rise to a moral hazard problem. To overcome this difficulty, leaders can use a very broad set of instruments to monitor their workers, such as procedures.

As a consequence, employers sometimes seem to care more about compliance than the final quality of the delivered service. This takes a dramatic extend in the medical sector where doctors spend two thirds of their time filling forms[2] rather than taking care of their patients. Bernstein (2012) empirically observes in a cellphone factory in China that better control of the agent's action deters productive behaviors. Bernstein calls this phenomenon the "transparency paradox" and shows that when an input measure is introduced by the principal, the agent is encouraged to follow the rules and not to do his best. These considerations have recently been highlighted by anthropologists such as David Graeber in his pamphlet "On the Phenomenon of Bullshit Jobs" published in *Strike! Magazine* in 2013 or sociologists as Hibou (2012). They show that workers, especially in large organizations, consider a large part of their job to be useless.

The contribution of this paper is to study the compatibility between a bureaucratic management of the agent and high effort on quality. Weber (1921) originally characterizes a bureaucratic management by the uniformity of the jobs and normalization of the tasks which are ensured through rules, codes and precise procedures. We define thus the procedure as the association of codification and verification. Consider, for instance, a clinical exam. The doctor has to perform the same routine over and over. Each step is well defined and the observations have to be duly noted in the patient's chart. Therefore, in order to control the doctor's effort, the principal can check the chart. A clinical exam is thus a procedural activity.

We consider a principal agent model with moral hazard, limited liability and where protagonists are risk neutral. The agent's actions lead to three levels of quality : low, acceptable and high. He can either shirk, work or go the extra mile. By working, the agent meets the

[2] A 2016 study of the Annals of Internal Medecine on Allocation of Physician Time in Ambulatory Practice shows that for every hour spent with the patient physicians have to invest up to two additional hours in paperwork.

standards of her profession and decreases the occurrence of bad performance[3]. A doctor will perform examination, follow the steps of diagnostic and treat the patient. By running the extra mile the agent increases the occurrence of the best performance. A doctor will go to the bedside of his patient at night to bring comfort and pep talk. We claim that some aspects of the effort on quality can not be codified, typically the ones associated to the extra mile. Hence, a procedure makes it easier for the principal to control only that at least the standard effort is exerted.

We show that such a procedure improves incentives because the principal can condition the payment of the agent to both an output signal - observed quality - and a fulfilled procedure. Codification and verification reduces the implementation costs no matter what the target selected by the principal is - acceptable quality or high quality. Nevertheless, the incentive improvements are stronger when only standard effort is implemented, encouraging the principal to disregard high quality. Therefore, there is an incompatibility between procedure and high effort on quality.

The principal chooses to introduce a procedure if available because implementation costs decrease no matter what the selected objective is. In an extended version of the model, with many clients to serve and a capacity constraint on the agent's side[4], we show that enforcing a procedure may not be optimal for the principal. The organization head aims at relaxing the limited liability constraint to reduce the agent's rent and thus entrusts his employee with many tasks. Since the agent is capacity constrained, she must give up some clients in order to perform the procedure.

Hence, when choosing whether to introduce procedure or not, the principal has to make a choice between two incentive strategies. In the first one, the agent has many clients to serve and is procedure free. In the second one, the agent treats fewer clients and is subject to procedure. We show that the first incentive strategy is more likely to be selected when the impact of the agent's effort on the probability of success is high, verification accuracy is low and time needed to perform procedure is high.

Procedures can be observed in numerous economical fields, especially in large organizations, and have an important role in the design of the incentives. Surprisingly, this feature is

[3]Note that a certain level of quality can be ensured in the medium effort. In a call center, for instance, some level of courtesy is guaranteed by the script used by the operators.

[4]We introduce capacity constraint on the agent's side to take into account that by making the agent invest in procedural activities the principal gives up some of her employee's time

understudied. Strausz (2006) is the only attempt to tackle this issue which we are aware. He introduces, in a moral hazard framework with limited liability, paperwork as a monitoring device and shows that the principal disregards the agent's cost increase and explains why organization heads may demand too much paperwork. Our definition of a procedure partly based on verification, costly for the agent, is close to the notion of paperwork developed by Strausz. Moreover, we do observe, in our results, a social inefficiency of verification. However, our approach significantly differs. We do not focus on quantity but on quality issues, we show how a procedure impacts the implementation of the extra mile.

This third option requires to reconsider a standard assumption in agency frameworks, namely the monotone likelihood ratio property. When there are information asymmetries, the principal has to condition the incentives to signals on unobservable variables. The main idea of the monotone likelihood ratio property is that a high level of unobservable variable is more likely if a high level of output is observed or, as it is expressed and discussed by Milgrom (1981) in the moral hazard framework, "greater profit are evidence of greater effort from the agent". Thus, if the payment increases with the output, the agent is likely to respond with more effort. This assumption is of great importance in the principal-agent literature (see Laffont and Martimort (2009) for a survey), the theories of organizing (see Koh (1992) for a survey) and financing (see Innes (1993) for a survey). In our model, however, when exerting the extra mile, the agent performs the standard effort and some additional actions. The extra mile is then no better than the standard effort when it comes to avoiding the worst outcome. The role of the extra mile is to favor the highest outcome. Therefore, the highest likelihood ratio when acceptable quality is observed can not be associated to the extra mile and the likelihood ratios can not be monotonic in effort.

From a technical point of view, the main result of our work comes from features of the limited liability constraint. The limited liability constraint is introduced to take into account the fact that the agents cannot receive a negative bonus i.e. cannot be prohibitively punished in case of failure. This generates, with moral hazard, a rent left to the agent as it is originally and extensively studied by Innes (1990). To reduce the rent, the principal can bundle the tasks. By doing so, she can use a single reward for several activities and merges the limited liability constraints in a unique relaxed condition. Laux (2001) highlights this effect in a multitask framework. Furthermore, he shows that if the agent has to perform observable and unobservable efforts, the principal can use the rent left for unobservable actions to finance the observable ones. Even if procedure is not an independent task and conveys information about the agent's doings, we observe similar mechanisms. Since procedure does not contribute to the outcome, the principal can condition the payment of the agent to a good signal without directly paying for it. The cost of procedure, if low enough, is therefore covered by the rent left to the agent for his unobservable actions.

On the one hand, our definition of procedure follows on from discussions on the value of information about the agent's action in the moral hazard framework. The key question originally asked and answered by Harris and Raviv (1979) is to explain the widespread use of contracts in which the result of an imperfect monitoring process is used to determined the compensation of the agent. They give a characterization of the contract with monitoring where the agent is paid according to a prescribed schedule if her action is judged acceptable on the basis of monitoring outcome [5]. This schedule is interpreted in Mirrlees (1976) as an *instruction*, "a promise that reward will vary rapidly in the neighborhood of a particular output level". The procedure defined in our model conveys similar intuitions and can be seen as an *instruction*.

On the other hand, our definition of procedure as a codification and a verification also echos in the multitasking literature. The issue in these frameworks is that incentives both motivate effort and determine its allocation among tasks weakening the incentive power (Holmstrom and Milgrom (1991), Dixit (1997)). In such frameworks, monitoring has a major role since it makes it possible to change the balance between the tasks. Sinclair-Desgagné (1999) shows that stronger incentives can be ensured through a selective audit scheme in which appreciation of less tangible actions is contingent to the observation of high performance achievements in more visible tasks. Our model shares some of these intuitions since the agent's remuneration is conditional on actions conform to an easy to verify codification. Yet, our results do not come from this intern mechanism of the multitasking framework since the procedure does not verify that the agent exerted the standard effort but at least the standard effort.

Our observations on organizational design are consistent with the analyses of the trade-off between monitoring and incentives proposed by Demougin and Fluet (2001) in a principal

[5] The main result of Harris and Raviv (1979) is to show that the remuneration of a risk averse agent will always partly depend on the information correlated to her effort. This result is independently demonstrated by Holmström (1979)

agent model with limited liability where the principal uses less monitoring when its costs increase. Moreover, these results echos with the findings of Zhao (2008) who shows in a multitask framework that using perfect monitoring only on a subset of tasks may weaken the incentives. However, it significantly differs from other theoretical traditions found in Baker (2002) or Baker et al. (1994). The issue here is not a misalignment between input measure and the output.

The rest of the paper proceeds as follows: Section 2 presents the optimal incentive schemes with the extra mile. In section 3 procedure is introduced and the main results are presented. Section 4 provides an extension with many clients to serve and capacity constraint on the agent's side while section 5 concludes. Proofs can be found in the Appendix.

2 A model of the extra mile

A risk neutral principal (she) hires a risk neutral agent (he) to perform a job. The output is a level of quality q either Low (L), Acceptable (A) or High (H). The principal values the quality q and $H > A > L = 0$ without lost of generality. The agent provides a non observable effort on quality $e \in \{0, 1, 2\}$ and faces a cost $c_e \in \{c_0, c_1, c_2\}$ where $c_2 > c_1 > c_0$. We assume that $c_0 = 0$ without lost of generality. Those efforts affect the probabilities of the output. $P_e(q)$ is the probability of getting the level of quality q knowing that effort e was exerted by the agent. The main idea of this extra mile model is that exerting at least a standard effort decreases the probability of getting the worst outcome. Exerting at least e=1 shifts the probability weights from L to $A \cup H$. Once this shift is operated, running the extra mile allows to reallocate some weight from A to H. Running the extra mile does not impact the probability of getting the worst but increases the probability of getting the highest quality. In order to capture this implicit sequentially of the extra mile we make two assumptions.

Assumption 1 (A1)

$$\frac{P_2(H)}{1 - P_2(L)} > \frac{P_1(H)}{1 - P_1(L)} = \frac{P_0(H)}{1 - P_0(L)}$$

This first assumption ensures that the probability of getting H, once at least a quality A is achieved, is higher when the highest effort is exerted and that this conditional probability is identical when efforts $e = 1$ or $e = 0$ are selected. The idea of this assumption is that the probability of getting H from A rises if and only if the extra mile is ran by the agent. Note that these ratios are not hazard rates, they are the probability of getting the best once the worst is avoided.

Assumption 2 (A2)

$$P_2(L) = P_1(L) < P_0(L)$$

This second assumption ensures that the probability of getting the worst is higher when the agent shirks and that the probability of avoiding the worst is identical with the highest efforts. The meaning of this assumption relies upon the extra mile logic again. The purpose of the standard effort - $e = 1$ - is to decrease the probability of getting the worst outcome. An agent who runs the extra mile does the same things as an agent who exerts the standard effort and some extra actions that increase the probability of getting the best outcome. Therefore, the probability to obtain L is identical with $e = 1$ and $e = 2$. Note that this assumption rules out cases where the highest effort increases both the probabilities of getting the best and the worst i.e. where running the extra mile is a risky action.

On the one hand, the extra mile implies unambiguously First Order Stochastic Dominance (see the appendix) but does not verify the Monotone Likelihood Ratio property (see the appendix). The extra mile naturally places the analysis in the particular case where FOSD and MLRP are not simultaneously verified. This come from the fact that the highest probability of getting exactly the acceptable quality is achieved when the standard effort is exerted. Therefore, the likelihood ratios can not be monotonic in effort. The probabilities are ordered in the following table[6].

	e=0		e=1		e=2
L	$P_0(L)$	$>$	$P_1(L)$	$=$	$P_2(L)$
A	$P_0(A)$	$<$	$P_1(A)$	$>$	$P_2(A)$
H	$P_0(H)$	$<$	$P_1(H)$	$<$	$P_2(H)$

[6]The nine probabilities $P_e(q)$ are variables. By definition, we know that $P_e(L)+P_e(A)+P_e(H) = 1 \ \forall e \in \{0, 1, 2\}$. At this point, there are six degrees of freedom remaining. The two assumptions we made impose order on these probabilities but only one equality $P_1(L) = P_2(L)$. Therefore, we lose only one degree of freedom with our assumptions

On the other hand, these two assumptions ensure that the highest likelihood ratio when the principal wants to implement $e = 2$ is the one associated with the output H and that the highest likelihood ratio when the principal wants to implement $e = 1$ is the one associated with A. This implication is of great importance here since in a moral hazard framework with limited liability and risk neutral protagonists it is optimal for the principal to pay only for one signal, the one associated to the highest likelihood ratio as it is pointed out in Fleckinger (2012) and displayed in **L1**.

Lemma 1 (L1) *The only optimally strictly positive wage in a moral hazard framework with limited liability and risk neutrality is the one associated to the highest likelihood ratio.*

The only available signal to the principal is the observed quality $q \in \{L, A, H\}$. Therefore the incentive scheme is a payments vector $W = (w_H, w_A, w_L)$ where the value w_q is the bonus paid to the agent when the principal observes quality q. The expected utility of the Agent $U(W, e)$ is the difference between expected remuneration and costs of efforts.

$$U(W, e) = P_e(L)w_L + P_e(A)w_A + P_e(H)w_H - c_e$$

And the expected profit of the principal is $V(W, e)$.

$$V(W, e) = P_e(A)(A - w_A) + P_e(H)(H - w_H)$$

The purpose of the remaining of this section is to study the trade-off faced by the principal between acceptable and high quality. In order to implement an high effort on quality, the principal has to find the minimum positive values of w_L, w_A and w_H such that it is in the best interest of the agent to exert $e = 2$. It is then optimal for the principal from **L1** to set $w_L^* = w_A^* = 0$ and $w_H^* > 0$. This optimal contract is noted W_2. There are two possible values for the optimal wage when the principal implements the extra mile depending on which incentive constraint is binding. To identify the best deviation of the agent with respect to the value of the parameters, we have to introduce here a measure of the relative convexity of the gains noted γ and defined as

$$\frac{P_2(H) - P_1(H)}{P_1(H) - P_0(H)} \equiv \gamma$$

This parameter will be crucial in our analyses. If the convexity of the benefits γ between $e = 2$ and $e = 1$ relative to $e = 1$ and $e = 0$ is higher than the relative convexity of the corresponding costs, then the best deviation for the agent is to shirk. Otherwise, the best deviation for the agent is to exert the standard effort. If the principal implements $e = 1$ the highest likelihood ratio, no matter what is the considered deviation, is associated to the output A. Therefore, the only strictly positive wage is w_A^* and the optimal contract is noted W_1. If the principal do not want the agent to work it is optimal to choose null payments and the optimal contract is noted W_0. These optimal contracts are given in **Proposition 1**.

Proposition 1 ***Implementation***

$W_2 = (w_H^*, 0, 0)$ *with*

$$w_H^* = \begin{cases} \frac{c_2 - c_0}{P_2(H) - P_0(H)} & \text{if } \frac{c_2 - c_1}{c_1 - c_0} \leq \frac{P_2(H) - P_1(H)}{P_1(H) - P_0(H)} \equiv \gamma \\ \\ \frac{c_2 - c_1}{P_2(H) - P_1(H)} & \text{otherwise} \end{cases}$$

$W_1 = (0, w_A^*, 0)$ *with*

$$w_A^* = \frac{c_1}{P_1(A) - P_0(A)}$$

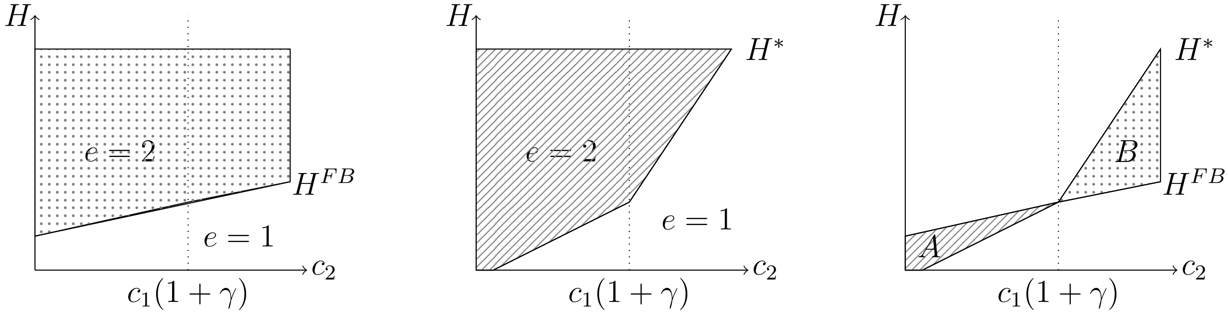
$W_0 = (0, 0, 0)$

The principal implements the extra mile if and only if H is high enough with respect to the values of the other parameters. The principal compares the current value of H to a threshold noted H^*. If $H \geq H^*$, the extra mile is implemented. In order to interpret this comparison, let us consider the corresponding threshold in a first best scenario where the principal can observe the agent's effort. This threshold is noted H^{FB}. The comparison between these thresholds leads to a first result : when c_2 is low enough - $c_2 \leq c_1(1 + \gamma)$ - the principal in second best chooses too much quality and too little when c_2 gets bigger compared to the first best case.

Proposition 2 ***Comparison***

$$\begin{cases} H^{FB} \geq H^* & if\ \frac{c_2-c_1}{c_1-c_0} \leq \gamma \\ \\ H^{FB} < H^* & otherwise \end{cases}$$

In the first best case, the trade off between high and acceptable quality is a comparison between expected gain and costs illustrated in figure (a). If the value of H gets higher or the value of c_2 gets smaller the principal is more likely to implement the extra mile. In the second best scenario, illustrated in figure (b), the principal deals with expected implementation costs. When $c_2 \leq c_1(1+\gamma)$, the best deviation for the agent is to shirk. This deviation, for those values of c_2, is relatively cheaper to correct for the principal and she implements high quality for lower values of H. This case corresponds to the hatched area of figure (c) labelled A. When c_2 gets bigger it is the deviation in $e = 1$ which is easier to correct and the second best threshold drifts away from the first best one, the principal needs higher value of H to implement the extra mile. This case corresponds to the dotted area of figure (c) labelled B.



(a) Threshold in First Best (b) Threshold in Second Best (c) Comparing H^{FB} and H^*

3 Introducing the procedure

We define the procedure as the association of codification and verification. This definition echoes with the seminal characterization of bureaucratic management by Weber (1921) where normalization of jobs is ensured through written rules, a systematic certification of the paperwork by the hierarchy and archiving. In the medical field, for instance, a clinical exam involves well codified thirty-eight steps each of them associated with a line to fill in the patient's chart. If the form is completed, the head of department assumes that the doctor followed the rules independently of the final outcome.

In a moral hazard context, codification is important since it pins down precisely what the principal expects of her agent. As a consequence, codification allows a more accurate verification of the agent's effort on this part of the job. This verification consists in "ticking boxes" and can take various forms: control panels in production chains, checking the paperwork, etc. Note that procedure does not impact the output: paperwork does not help the doctor to perform the examination or to find the good medication, it is just informative about the agent's action on the codified part of his job.

Some parts of the job are yet more complex and hardly convertible into rules. We claim this is the case of the additional effort on quality - running the extra mile. The purpose of the current section is to study the modification of the trade off faced by the principal between acceptable and high quality when she is able to codify and verify a part of the agent's effort. In our model, effort $e = 1$ can be codified and partially verified.

When the procedure is in place, an additional signal σ is generated which is either good, the agent followed the rules ($\sigma = 1$), or bad ($\sigma = 0$). Moreover, filling forms takes time and induces a cost c_f supported by the agent. The procedure imperfectly identifies whether at least a standard effort is exerted. The informational structure is given in the following table.

$Prob(\sigma = i \mid e = j)$	e=0	e=1	e=2
$\sigma = 0$	m	0	0
$\sigma = 1$	$1 - m$	1	1

An agent who exerts at least the standard effort $e = 1$ can not send a bad signal through the procedure, and a non compliant agent is spotted with probability $m \in [0, 1]$. We define m as the quality of the procedure [7]. In our model, all the actions constitutive of the effort $e = 1$ are assumed to be codifiable. This means that with or without codification an agent

[7] This informational structure is "hostile" in the terminology of Fleckinger et al. (2017). That means

who exerts at least $e = 1$ carries out the same tasks in the same way. As a consequence, when a coherent procedure is introduced it is impossible that a compliant agent send a bad signal as long as he pays the cost of procedure. If the agent shirks he has to falsify his report making mistakes with probability $m \in [0, 1]$.

Therefore, in addition to the final outcome, the principal receives a signal through the procedure and the payments vector is $W_f = (w_{L,0}, w_{A,0}, w_{H,0}, w_{L,1}, w_{A,1}, w_{H,1})$. The possible wages are then conditional to both final quality and signal of the procedure. If at least a standard effort is implemented by the principal, it is optimal to pay the agent only when a good signal occurs along with the relevant output signal. Thereby, the only strictly positive wage when $e = 2$ and $e = 1$ are implemented are respectively $w^*_{H,1}$ and $w^*_{A,1}$ (see appendix) and are lower that without the procedure. An implication of the introduction of the procedure is that the participation constraint could be binding. This happens when the cost c_f is higher than the agent's expected benefit of complying. In this case, the principal continues to pay the agent the optimal bonus in order to satisfy the incentive constraints and has to compensate him for the difference between the cost of procedure and his gains from the incentive dimension. In the following, we will focus our analyse on cases where the participation constraint is not binding: our aim is not to study when the agent suffers procedure but its impact on the incentive dimension.

Proposition 3 ***Implementation with a procedure***

If the participation constraint is not binding with the procedure, the principal implements the same effort as without, at a lower wage cost.

What is striking is that the principal does not directly pay for this monitoring strategy if c_f is small enough[8]. The principal continues to condition the payment to the output - given the fact that procedure can't be a failure if at least $e = 1$ is exerted - and finances the procedure with agent's rent which has a negative impact on overall welfare as it is stated in Corollary 1.

that a deviation of the agent can be identified for sure only when a bad signal occurs. Other informational structures can be considered. Think, for instance, of a "friendly" informational structure where the extra mile can be identified for sure only when a good signal occurs. In this case, procedure imperfectly identifies whether $e = 2$ is exerted. This informational structure seems irrelevant since the best effort is precisely the hardest to observe.

[8]The case where the participation constraint is binding is discussed in a closely related framework by Strausz (2006)

Corollary 1 *If the participation constraint is not binding, the procedure reduces welfare.*

This result shares some intuitions with Strausz (2006) who shows, in a closely related model, that internal paperwork allows more accurate monitoring and in her decision the principal disregards the agent's cost increase.

The optimal wages with the procedure are very similar to the ones described in the case without procedure. The difference stands on the m parameter and comes from the fact that a single wage is used to pay for two activities. In a moral hazard framework with limited liability and risk neutral protagonists Laux (2001) shows that such bundling makes it possible to reduce the rent left to the agent since it relaxes the limited liability constraint. We do observe similar reduction once the procedure is introduced. The difference of the wages when $e = 1$ is implemented is

$$w_A^* - w_A^{1*} = \frac{c_1 m P_0(A)}{[P_1(A) - P_0(A)][P_1(A) - (1-m)P_0(A)]}$$

This difference is positive and increasing in m. The corresponding difference when the extra mile is implemented has to be considered with respect to the value of c_2 and has the same properties: the procedure never increases the implementation costs. The introduction of the procedure has an effect on the binding constraints when $e = 2$ is implemented. The interval where $e = 0$ is the best deviation shrinks.

$$\gamma = \gamma(0) \geq \gamma(m)$$

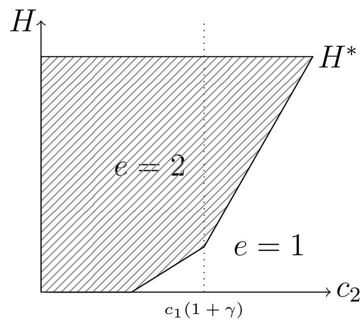
The distance between those values increases with m. This comes from the fact that with the procedure the principal can easily verified that at least $e = 1$ is exerted. Therefore, the agent is more likely to deviate in $e = 1$. Considering the difference of wages with and without the procedure, three cases has to be considered.

For small values of c_2, procedure induces increasing gains in m. For intermediary values of c_2 the principal has to correct a deviation in $e = 0$ without the procedure and in $e = 1$ with the procedure. For those values of c_2 the former implies lower compensation and the difference of implementation costs is positive. When c_2 is high, the principal faces the same incentive problem with and without the procedure and there is no difference in the implementation costs. The impact of the introduction of the procedure on the trade off between high and acceptable quality is given in **Proposition 4**.

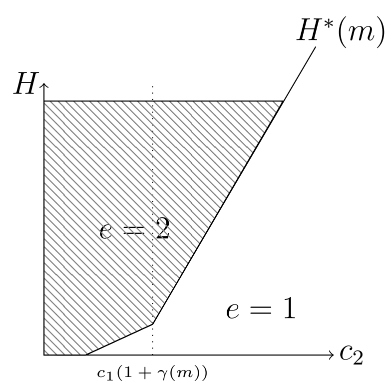
Proposition 4 ***Comparison with the procedure***
The introduction of the procedure makes high effort on quality both cheaper and less implemented than without the procedure.

The conditions under which the principal ask her agent to run the extra mile, when the procedure is introduced, is again a threshold of H noted $H^*(m)$ represented in figure (f). Comparing this new threshold with the one of **Proposition 2**, drawn independently in figure (d), we see that the space where high quality is implemented shrinks when procedure is introduced. The couples of H and c_2 for which the principal implements the extra mile without the procedure but only the standard effort with the procedure is represented by the hatched area labelled D. This comes from the fact that by choosing $e = 2$ the principal gives up some incentives gains generated by the introduction of procedure when only acceptable quality is implemented. Therefore the principal needs higher values of H to find high quality attractive.

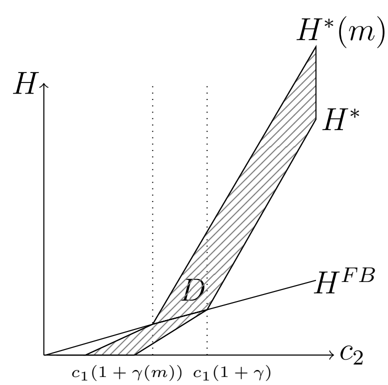
This result significantly differs from the mechanisms identified in the multitasking literature. Indeed, the introduction of the procedure allows to partially verify that at least a standard effort is exerted. Thereby, this informational structure does not directly favor the implementation of standard effort. The modification of the principal's choice with the procedure comes from the extra mile nature of the production function. The ability to codify and verify a part of the agent's job discourages the principal to ask for more.



(d) Threshold without procedure



(e) Threshold with procedure



(f) Comparing H^* and $H^*(m)$

The introduction of the procedure has an ambiguous effect on the second best threshold compared to the first best one. On the one hand, for low values of c_2, the procedure implies a rotation toward the First Best. This comes from the fact that, when the procedure is introduced, shirking is detected more easily. As a consequence, the deviation in $e = 0$ becomes less profitable for the agent when the extra mile is implemented and cheaper to correct when the standard effort is implemented. In the other hand, for higher values of c_2, the introduction of procedure pushes the threshold away from its First Best position due to the reduction of the implementation cost of standard effort with the procedure.

4 Task Bundling vs. the procedure

In the framework described in section 2 and 3, procedure is always introduced if available no matter what is the targeted level of quality. Indeed, when acceptable quality is implemented a procedure increases the highest likelihood ratio and reduces the expected implementation costs.

$$\frac{P_1(A)}{P_0(A)} \leq \frac{P_1(A)}{(1-m)P_0(A)}$$

Similarly, a procedure never strictly decreases the incentive efficiency when the extra mile is implemented.

$$\begin{cases} \frac{P_2(H)}{P_0(H)} \leq \frac{P_2(H)}{(1-m)P_0(H)} & \text{if } \frac{c_2-c_1}{c_1-c_0} \leq \gamma \\ \\ \frac{P_2(H)}{P_1(H)} = \frac{P_2(H)}{P_1(H)} & \text{otherwise} \end{cases}$$

However, procedure seems to be more frequent in particular cases such as large organizations and administrations. There are two possible explanations for this phenomenon. On the one hand, simple and recurrent tasks are more likely to appear in those institutions. On the other hand, the codification and verification implied by a procedure does not convey the same incentive power for all kinds of jobs.

Consider a case where the principal has many clients to serve and a single agent can not provide for everyone. Indeed, there is no reason to think that someone who is asked to produce high quality service and write a substantial amount of paperwork can treat as many clients as an agent who only has to ensure acceptable quality.

Taking into account this observation we impose a capacity constraint on the agent's side. We assume that each agent has only a period of length T to spare. In this setup, the principal may hire as many employees as she wants. All types of activity use some of the agent's time resource. Thus, providing standard effort for a client takes a time t_1, t_2 is the time needed to perform the extra mile where $t_1 < t_2$ and fulfilling procedure implies an investment t_f. In this configuration, a deviation gives free time to the agent [9].

The amount of clients for each agent depends on the quality targeted by the principal and the introduction of the procedure. The principal when she chooses an effort and whether induce procedure or not also chooses a type for her incentive scheme. For instance, when high quality is selected an agent may only take care of $n_2 = T/t_2$ agents. When only acceptable quality is expected the same agent will be in charge of $n_1 = T/t_1$ clients. Obviously $n_1 > n_2$[10].

Procedure does not always lower the implementation costs in this new framework. Inducing procedure diminishes the number of projects carried out by an agent and thus may weaken the incentives. Thereby, the principal faces a trade-off when choosing to introduce procedure or not between tasks bundling and gains from codification and verification. The determinants of this choice when a standard effort is implemented are given in **Proposition 5**[11].

Proposition 5 *With many clients to serve and a capacity constraint on the agent's side, the average expected wage decreases when the procedure is introduced and standard effort is implemented if and only if*

$$(1-m) \leq \left[\frac{P_0(A)}{P_1(A)}\right]^{\frac{t_f}{t_1}}$$

This means that the principal is more likely to choose procedure when filling forms is not too time consuming. She will also select procedure when the joint measure of the wage's

[9]Note that, in this model, we only consider the time needed to perform the job, not the intensity of the effort. Further discussions on this issue could be engaged but are not central for the comments we make on organizational design based on the trade off between task bundling and procedure.

[10]Similar observation can be done on procedure. The number of clients treatable by a single agent when standard effort is implemented and procedure introduced is $n_{1,f} = T/(t_1 + t_f)$ while the number when high quality is ordered is $n_{2,f} = T/(t_2 + t_f)$

[11]The corresponding observation when extra mile is implemented can be found in the appendix

sensitivity to effort is bigger than the probability to believe the agent's falsification. That means that when principal may encounter dramatic situation, cases where she has to pay huge bonuses, she will rather pick procedure which appears as a "safe" incentive policy and implies a larger number of agents. It is then not surprising to find intensive use of procedure - and then large institutions - in constrained budget sectors (administration, tertiary activities, etc.). Moreover, the principal chooses procedure if the impact of the effort on success is sufficiently low. This result is consistent with the considerations developed by Wilson (1989) in his famous work on bureaucracy. He identifies some activities such as teaching and law enforcement where the agent's behaviour is the key determinant of success making the introduction of rules difficult and sometimes counterproductive.

5 Conclusive remarks

In a moral hazard setup with limited liability the principal can not appropriate all rents. This issue is magnified when the task performed by the agent is complex. To monitor the agent's action, organization heads can use several intruments such as procedure.

We model procedure as a nonproductive task that is informative about the agent's doing. Conditioning the payment on a well filled form partially relaxes the limited liability constraint and reduce the implementation costs. Moreover, since procedure does not directly play a role in productivity the principal is not compelled to explicitly pay for it and rely upon the rent left to the agent for unobservable actions at society's expense.

When the highest quality depends on exceptional dedication of the agent we show that a procedure introduced over a part of the agent's job makes the extra mile both cheaper to enforce and less implemented than without the procedure. This comes from the fact that the incentive gains of verification are higher when only standard effort is implemented. Therefore, a better control of the agent lowers quality.

In the basic framework, using a procedure is always optimal for the principal. In an extension with many clients to serve and capacity constraint on the agent's side we show that the principal faces a trade-off when choosing whether to introduce a monitoring strategy or not. The principal has to choose between two types of organization. One based on a small number of employees in charge of an important amount of client, free of procedure and more likely to go the extra mile. The other with more agents subject to procdure and less likely to go the extra mile.

References

Baker, G. (2002). Distortion and risk in optimal incentive contracts. *Journal of human resources*, pages 728–751.

Baker, G., Gibbons, R., and Murphy, K. (1994). Subjective performance measures in optimal incentive contracts. *The Quarterly Journal of Economics*.

Bernstein, E. S. (2012). The transparency paradox a role for privacy in organizational learning and operational control. *Administrative Science Quarterly*, 57(2):181–216.

Demougin, D. and Fluet, C. (2001). Monitoring versus incentives. *European Economic Review*, 45(9):1741–1764.

Dixit, A. (1997). Power of incentives in private versus public organizations. *The American Economic Review*, 87(2):378–382.

Fleckinger, P. (2012). Correlation and relative performance evaluation. *Journal of Economic Theory*, 147(1):93–117.

Fleckinger, P., Glachant, M., and Moineville, G. (2017). Incentives for quality in friendly and hostile informational environments. *American Economic Journal: Microeconomics*, 9(1):242–274.

Harris, M. and Raviv, A. (1979). Optimal incentive contracts with imperfect information. *Journal of economic theory*, 20(2):231–259.

Hibou, B. (2012). *La bureaucratisation du monde à l'ère néolibérale*. La Découverte.

Holmstrom, B. and Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, & Organization*, 7:24–52.

Holmström, B. (1979). Moral hazard and observability. *Bell Journal of Economics*.

Innes, R. (1993). Financial contracting under risk neutrality, limited liability and ex ante asymmetric information. *Economica*, pages 27–40.

Innes, R. D. (1990). Limited liability and incentive contracting with ex-ante action choices. *Journal of economic theory*, 52(1):45–67.

Koh, W. T. (1992). A note on modelling tournaments. *Journal of Economics*, 55(3):297–308.

Laffont, J.-J. and Martimort, D. (2009). *The theory of incentives: the principal-agent model*. Princeton university press.

Laux, C. (2001). Limited-liability and incentive contracting with multiple projects. *RAND Journal of Economics*, pages 514–526.

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, pages 380–391.

Mirrlees, J. A. (1976). The optimal structure of incentives and authority within an organization. *The Bell Journal of Economics*, pages 105–131.

Sinclair-Desgagné, B. (1999). How to restore higher-powered incentives in mulitask agencies. *The Journal of Law, Economics and Orgarnization*.

Strausz, R. (2006). Buried in paperwork: Excessive reporting in organizations. *Journal of Economic Behavior & Organization*, 60(4):460–470.

Weber, M. (1978 (1921)). *Economy and society: An outline of interpretive sociology*. Univ of California Press.

Wilson, J. Q. (1989). *Bureaucracy: What government agencies do and why they do it*. Basic Books.

Zhao, R. R. (2008). All-or-nothing monitoring. *The American Economic Review*, 98(4):1619–1628.

6 Appendix

A1 and A2 imply First Order Stochastic Dominance

By definition,

$$P_0(L) + P_0(A) + P_0(H) = P_1(L) + P_1(A) + P_1(H) = P_2(L) + P_2(A) + P_2(H) = 1$$

A2 gives

$$P_0(L) \geq P_1(L) \geq P_2(L)$$

From A2 we get $P_2(L) = P_1(L)$ and from A1 we get $\frac{P_2(H)}{1-P_2(L)} > \frac{P_1(H)}{1-P_1(L)}$. These lead to

$$P_1(L) + P_1(A) \geq P_2(L) + P_2(A)$$

From A2 we get $P_0(L) > P_1(L)$ and from A1 we get $\frac{P_0(H)}{1-P_0(L)} = \frac{P_1(H)}{1-P_1(L)}$. These lead to

$$P_0(L) + P_0(A) \geq P_1(L) + P_1(A)$$

From the last two inequalities we obtain

$$P_0(L) + P_0(A) \geq P_1(L) + P_1(A) \geq P_2(L) + P_2(A)$$

And A1 and A2 imply First Order Stochastic Dominance

A1, A2 and the likelihood ratios

Consider a case where the principal wants to implement the effort e_2. From A1 we get

$$\frac{P_2(H)}{P_2(H)+P_2(A)} \geq \frac{P_1(H)}{P_1(H)+P_1(A)} \iff \frac{P_2(H)}{P1(H)} \geq \frac{P_2(A)}{P1(A)}$$

And,

$$\frac{P_2(H)}{P_2(H)+P_2(A)} \geq \frac{P_0(H)}{P_0(H)+P_0(A)} \iff \frac{P_2(H)}{P_0(H)} \geq \frac{P_2(A)}{P_0(A)}$$

Since A1 and A2 imply First Order Stochastic Dominance we get as well

$$\frac{P_2(H)}{P_1(H)} \geq \frac{P_2(L)}{P_1(L)}$$

And,

$$\frac{P_2(H)}{P_0(H)} \geq \frac{P_2(L)}{P_0(L)}$$

Consider now the case where the principal wants to implement the effort e_1. We have already shown that

$$\frac{P_1(A)}{P_2(A)} \geq \frac{P_1(H)}{P_2(H)}$$

First Order Stochastic Dominance ensures that,

$$\frac{P_1(A)}{P_0(A)} \geq \frac{P_1(L)}{P_0(L)}$$

A1 and A2 implies that

$$P_1(A) \geq P_2(A) \iff \frac{P_1(A)}{P_2(A)} \geq \frac{P_1(L)}{P_2(L)} = 1$$

Finally we deduce that

$$\frac{P_1(A)}{P_0(A)} \geq \frac{P_1(H)}{P_0(H)}$$

Proof of Proposition 1

(a) When $e = 2$ is implemented by the principal she faces participation, incentive and limited liability constraints (outside option has no value in our model). Since our framework is compatible with the findings in Innes (1990) the only strictly positive wage is the one associated with the highest likelihood ratio - w_H. This can be demonstrated as well with incentive efficiency ratio (Fleckinger (2012)). The remaining incentive constraints are

$$P_2(H)w_H - c_2 \geq P_1(H)w_H - c_1$$
$$P_2(H)w_H - c_2 \geq P_0(H)w_H$$

Finally we get ,

$$w_H^* = \begin{cases} \frac{c_2 - c_0}{P_2(H) - P_0(H)} & \text{if } \frac{c_2 - c_1}{c_1 - c_0} \leq \frac{P_2(H) - P_1(H)}{P_1(H) - P_0(H)} \\ \\ \frac{c_2 - c_1}{P_2(H) - P_1(H)} & \text{otherwise} \end{cases}$$

(b) When the principal implements $e = 1$, the highest likelihood ration o matter what is the considered deviation is associated with signal A. Therefore the only strictly positive wage is w_A. The optimal value of this wage is the solution of the following incentive constraint.

$$P_1(A)w_A - c_1 = P_0(A)w_A$$

The deviation in $e = 2$ is always dominated since $P_1(A) > P_2(A)$ and $c_2 > c_1$. Finally, we get

$$w_A^* = \frac{c_1}{P_1(A) - P_0(A)}$$

Proof of Proposition 2

(a) In the first best case where the principal observes the agent's action, her expected utility when the extra mile is implemented is

$$V(c_2, c_2, c_2, 2) = P_2(H)H + P_2(A)A - c_2$$

The corresponding expected utility when $e = 1$ is implemented is

$$V(c_1, c_1, c_1, 1) = P_1(H)H + P_1(A)A - c_1$$

Therefore, $V(c_2, c_2, c_2, 2) \geq V(c_1, c_1; c_1, 1)$ if and only if

$$H \geq \frac{A[P_1(A) - P_2(A)]}{P_2(H) - P_1(H)} + \frac{c_2}{P_2(H) - P_1(H)} - \frac{c_1}{P_2(H) - P_1(H)} \equiv H^{FB}$$

(b) In the second best case, the expected utility of the principal when the extra mile is implemented is

$$V(w_H^*, 0, 0, 2) = P_2(H)(H - w_H^*) + P_2(A)A$$

The corresponding expected utility when $e = 1$ is implemented is

$$V(0, w_A^*, 0, 1) = P_1(H)H + P_1(A)(A - w_A^*)$$

Therefore, $V(w_H^*, 0, 0, 2) \geq V(0, w_A^*, 0, 1)$ if and only if

$$H \geq \frac{A[P_1(A) - P_2(A)]}{P_2(H) - P_1(H)} + \frac{c_2}{P_2(H) - P_1(H)} - \frac{c_1}{P_2(H) - P_1(H)} \equiv H^{FB}$$

(c) We have now to study to relative position of the thresholds H^{FB} and H^*. Consider the extreme case where $c_2 = 0$. In this case H^{FB} is always higher :

$$c_1 \leq \frac{P_1(A)c_1}{P_1(A) - P_0(A)} \iff P_1(A) > P_1(A) - P_0(A)$$

Consider now the case where $\frac{c_2-c_1}{c_1-c_0} = \frac{P_2(H)-P_1(H)}{P_1(H)-P_0(H)}$. For this value of c_2 the thresholds are equal due to A1.

$$\begin{aligned}\frac{P_2(H)c_2}{P_2(H)-P_0(H)} - \frac{P_1(A)c_1}{P_1(A)-P_0(A)} = c_2 - c_1 &\iff \frac{P_0(A)}{P_0(H)} = \frac{P_1(A)}{P_1(H)} \\ &\iff \frac{P_1(H)}{P_1(A)+P_1(H)} = \frac{P_0(H)}{P_0(A)+P_0(H)}\end{aligned}$$

For higher values of c_2, the slope of H^* is always higher.

$$\frac{P_2(H)c_2}{P_2(H)-P_1(H) > c_2} \iff P_2(H) > P_2(H) - P_1(H)$$

Therefore, H^{FB} is higher than H^* if c_2 is small enough.

$$\begin{cases} H^{FB} \geq H^* & \text{if } \frac{c_2-c_1}{c_1-c_0} \leq \gamma \\ \\ H^{FB} < H^* & \text{otherwise} \end{cases}$$

The additional signal convied by procedure does not modify the order of the likelihood ratios

Consider a case where the principal wants to implement the effort e_2 and procedure is introduce. Let $P_{e,\sigma}(q)$ be the probability of observing quality q and signal σ when effort e is exerted. From section 2 we directly have

$$\frac{P_{2,1}(H)}{P_{1,1}(H)} = \frac{P_2(H)}{P_1(H)} \geq \frac{P_2(A)}{P_1(A)} = \frac{P_{2,1}(A)}{P_{1,1}(A)}$$

And,

$$\frac{P_2(H)}{P_0(H)} \geq \frac{P_2(A)}{P_0(A)} \iff \frac{P_2(H)}{(1-m)P_0(H)} \geq \frac{P_2(A)}{(1-m)P_0(A)}$$

We get as well,

$$\frac{P_{2,1}(H)}{P_{1,1}(H)} = \frac{P_2(H)}{P_1(H)} \geq \frac{P_2(L)}{P_1(L)} = \frac{P_{2,1}(L)}{P_{1,1}(L)}$$

And,

$$\frac{P_2(H)}{P_0(H)} \geq \frac{P_2(L)}{P_0(L)} \iff \frac{P_2(H)}{(1-m)P_0(H)} \geq \frac{P_2(L)}{(1-m)P_0(L)}$$

Consider now the case where the principal wants to implement the effort e_1 and a procedure is introduced. We directly have,

$$\frac{P_1(A)}{P_0(A)} \geq \frac{P_1(L)}{P_0(L)} \iff \frac{P_1(A)}{(1-m)P_0(A)} \geq \frac{P_1(L)}{(1-m)P_0(L)}$$

And

$$\frac{P_{1,1}(A)}{P_{2,1}(A)} = \frac{P_1(A)}{P_2(A)} \geq \frac{P_1(L)}{P_2(L)} = \frac{P_{1,1}(L)}{(P_{2,1}(L)} = 1$$

Finally we deduce that

$$\frac{P_1(A)}{P_0(A)} \geq \frac{P_1(H)}{P_0(H)} \iff \frac{P_1(A)}{(1-m)P_0(A)} \geq \frac{P_1(H)}{(1-m)P_0(H)}$$

Proof of Proposition 3

(a) When $e = 2$ is implemented and a procedure is introduced, the only strictly positive wage is the one associated with the highest likelihood ratio - $w_{H,1}$. The remaining constraints are

$$P_2(H)w_{H,1} - c_2 - c_f \geq P_1(H)w_{H,1} - c_1 - c_f$$
$$P_2(H)w_{H,1} - c_2 - c_f \geq (1-m)P_0(H)w_{H,1} - c_f$$
$$P_2(H)w_{H,1} - c_2 - c_f \geq 0$$

Note that when a procedure is introduced the agent faces an additional cost c_f. This cost has no consequence on the incentive constraints but implies that the participation constraint is no longer automatically satisfied. In fact, if $c_f > P_2(H)w_{H,1} - c_2$ the principal as to pay the difference between the rent left to the agent and the cost of the procedure.
Finally, when $c_f \leq P_2(H)w_{H,1} - c_2$, we get

$$w^*_{H,1} = \begin{cases} \frac{c_2 - c_0}{P_2(H)-(1-m)P_0(H)} & \text{if } \frac{c_2-c_1}{c_1-c_0} \leq \frac{P_2(H)-P_1(H)}{P_1(H)-(1-m)P_0(H)} \\ \\ \frac{c_2-c_1}{P_2(H)-P_1(H)} & \text{otherwise} \end{cases}$$

(b) The optimal strictly positive bonus when $e = 1$ is choose is found following the same path and, when the participation constraint is not binding

$$w^*_{A,1} = \frac{c_1}{P_1(A) - (1-m)P_0(A)}$$

Proof of Proposition 4

(a) The principal implements the extra mile with procedure if and only if $V(w^*_{H,1}, 0, 0, 2) \geq V(0, w^*_{A,1}, 0, 1)$.

$$H \geq \frac{A[P_1(A) - P_2(A)]}{P_2(H) - P_1(H)} + \frac{P_2(H)w^*_{H,1} - P_1(A)w^*_{A,1}}{P_2(H) - P_1(H)} \equiv H^*(m)$$

(b) Consider the extreme case where $c_2 = 0$. It is straightforward that the threshold with procedure is higher than the one without procedure.

$$\frac{P_1(A)c_1}{P_1(A) - P_0(A)} \geq \frac{P_1(A)c_1}{P_1(A) - (1-m)P_0(A)} \iff mP_0(A) \geq 0$$

Moreover, the threshold in second best with procedure is lower than the one of the first best

$$c_1 \geq \frac{P_1(A)}{P_1(A) - (1-m)P_0(A)} \iff P_1(A) \geq P_1(A) - (1-m)P_0(A)$$

The threshold of the first best and the second best with procedure are equal when $c_2 = c_1(1 + \gamma(m))$ due to A1 (cf. proof of proposition 5). Therefore when $c_2 \leq c_1(1 + \gamma(m))$ the threshold with procedure is above the one without procedure.

Finally, when $c_2 > c_1(1 + \gamma(m))$ the slope of the threshold with procedure is higher or equal to the one without procedure and the space where $e = 2$ is implemented shrinks when a procedure is introduced.

Proof of Proposition 5

When there are many clients to serve and a capacity constraint on the agent's side the average expected wage when acceptable quality is implemented decreases if and only if the highest likelihood ratio increases. Therefore, the introduction of a procedure reduces the implementation costs if and only if

$$\left[\frac{P_1(A)}{P_0(A)}\right]^{n_1} \leq \left[\frac{P_1(A)}{(1-m)P_0(A)}\right]^{n_{1,f}} \iff (1-m) \leq \left[\frac{P_0(A)}{P_1(A)}\right]^{\frac{t_f}{t_1}}$$

The introduction of the procedure has also an ambiguous effect when the extra mile is implemented. In particular, the measures of the relatives convexity of the gains with and without the procedure respectively $\gamma(m, n_{2,f})$ and $\gamma(0, n_2)$, defined below, have not a fixed order.

$$\begin{cases} \gamma(m, n_{2,f}) & \equiv \frac{P_2^{n_{2,f}}(H) - P_1^{n_{2,f}}(H)}{P_1^{n_{2,f}}(H) - (1-m)^{n_{2,f}} P_0^{n_{2,f}}(H)} \\ \\ \gamma(0, n_2) & \equiv \frac{P_2^{n_2}(H) - P_1^{n_2}(H)}{P_1^{n_2}(H) - P_0^{n_2}(H)} \end{cases}$$

If $\gamma(m, n_{2,f}) \leq \gamma(0, n_2)$, shirking is less likely to be selected by the agent. This can happen because reducing the number of tasks increases the cost of correcting deviation in $e = 1$ more than the cost to prevent shirking. In this configuration, the average expected wage decreases when the procedure is introduced and extra mile is implemented when $t_2 \leq t_1(1 + \gamma(m, n_{2,f}))$ if and only if

$$(1-m) \leq \left[\frac{P_0(H)}{P_2(H)}\right]^{\frac{t_f}{t_2}}$$

The corresponding condition when $t_1(1 + \gamma(m, n_{2,f})) < t_2 \leq t_1(1 + \gamma(0, n_2))$ is

$$\frac{t_2 - t_1}{t_1} \leq \frac{P_2^{n_2}(H)P_2^{n_{2,f}}(H) - P_2^{n_2}(H)P_1^{n_{2,f}}(H)}{P_2^{n_2}(H)P_1^{n_{2,f}}(H) - P_2^{n_{2,f}}(H)P_0^{n_2}(H)}$$

The average expected wage never decreases when a procedure is introduced and extra mile is implemented when $t_2 > t_1(1 + \gamma(0, n_2))$.

If $\gamma(m, n_{2,f}) \geq \gamma(0, n_2)$, shirking is less easy to prevent when the procedure is introduced. In this case introducing the provide is worth only for very low value of t_2. The average expected wage decreases when the procedure is introduced and extra mile is implemented when $t_2 \leq t_1(1+\gamma(0,n_2))$ if and only if

$$(1-m) \leq \left[\frac{P_0(H)}{P_2(H)}\right]^{\frac{t_f}{t_2}}$$

The average expected wage never decreases when a procedure is introduced and extra mile is implemented when $t_2 > t_1(1+\gamma(0,n_2))$.