Bounds on the Welfare Loss of Moral Hazard with Limited Liability

Felipe Balmaceda
*Universidad Diego Portales*

Santiago Balseiro
*Duke University*

José Correa
*Universidad de Chile*

Nicolas Stier-Moses
*Columbia University*

*Septiembre 2014*
Bounds on the Welfare Loss of Moral Hazard with Limited Liability

F. Balmaceda*, S.R. Balseiro**, J.R. Correa†, N.E. Stier-Moses‡

* Economics Department, Universidad Diego Portales, Chile  
felipe.balmaceda@udp.cl

** Fuqua School of Business, Duke University, USA  
srb43@duke.edu

† Industrial Engineering Department, Universidad de Chile, Chile  
jcorrea@dii.uchile.cl

‡ Graduate School of Business, Columbia University, USA  
School of Business, Universidad Torcuato Di Tella and CONICET, Argentina  
stier@gsb.columbia.edu

September 22, 2014

Abstract

This article studies a principal-agent problem with discrete outcome and effort level spaces. The principal and the agent are risk neutral and the latter is subject to limited liability. We consider the ratio between the first-best social welfare to the social welfare arising from the principal’s optimal pay-for-performance contract, i.e., the welfare loss. In the presence of moral hazard, we provide simple parametric bounds on the welfare loss of a given instance, and then study the worst-case welfare loss among all instances with a fixed number of effort and outcome levels. Key parameters to these bounds are the number of possible effort levels, the likelihood ratio evaluated at the highest outcome, and the ratio between costs of the highest and the lowest effort levels. As extensions, we also look at linear contracts and at cases with multiple identical tasks. Our work constitutes an initial effort to analyze losses arising from moral hazard problems when the agent is subject to limited liability, and shows that these losses can be costly in the worst case.

KEYWORDS: Principal-Agent Problem, Moral Hazard, Limited Liability, Welfare Loss, Price of Anarchy.
1 Introduction

The principal-agent model with moral hazard has been the workhorse paradigm to understand many interesting economic phenomena where incentives play a crucial role such as the theory of insurance under moral hazard (Spence and Zeckhauser, 1971), the theory of managerial firms (Alchian and Demsetz, 1972; Jensen and Meckling, 1979), optimal sharecropping contracts between landowners and tenants (Stiglitz, 1974), the efficiency wages theory (Shapiro and Stiglitz, 1984), financial contracting (Holmström and Tirole, 1997; Innes, 1990), and job design and multi-tasking (Holmström and Milgrom, 1991). Casual observation also suggests that moral hazard could be of practical importance. In fact, most sales workers are paid according to a fixed wage and either a bonus paid when a certain sales target is achieved or a commission rate over total sales. Franchisees are also motivated by contracts that entail a fixed payment and an agreement about how to share profits or sales. Additionally, managerial contracts usually consist of a combination of fixed wages and payments that are conditioned on performance. In short, incentive contracts are ubiquitous to the market economies.

Regardless of the reason for moral hazard, in most cases this entails welfare losses that remain as far as we know quantitatively uncounted for. This paper undertakes the task of quantifying the welfare losses implied by the existence of moral hazard in a principal-agent relationship with risk neutral individuals and limited liability.

The main consequences of moral hazard are by now well understood and deeply rooted in the economics of information literature, thus the moral-hazard paradigm is ripe for a deeper analysis of the quantitative, rather than qualitative, consequences of it. The setting we consider consists of a risk-neutral principal who hires a risk-neutral agent subject to limited liability to exert costly effort. The effort level and outcome space are discrete and bounded, and effort influences the distribution of output and cannot be observed by the principal. No restrictions are imposed on the slope of the contract. If the principal wishes to induce the agent to choose a given effort level, he should reward the agent when the realization of output is most indicative that the desired effort level has been chosen and he should punish him when a different outcome is observed. Because limited liability imposes a lower bound on the size of the punishment, the equilibrium contract leaves a limited liability rent to the agent. As a result the equilibrium contract might not maximize social welfare and the first-best outcome might not be attained; instead, the constrained contract will be second-best.\footnote{See, Balmaceda et al. (2012) for an extension to a continuous effort space. Bastin et al. (2013) generalize our results by considering assumptions that do not ensure the validity of the first-order approach.}

\footnote{When the participation constraint—rather than the limited-liability constraint—binds, providing incentives is costless since the agent cares only about the expected compensation and the participation constraint binds on his expected payoff. Thus, we focus on the case in which the parameters are such that this does not occur.}
1.1 Main Contributions

To quantify the inefficiencies introduced by moral hazard and limited liability, we measure the welfare loss introduced by a given contract, and rely on the concept of price of anarchy. The latter refers to the worst-case welfare loss in a non-cooperative game, that is, the welfare at equilibrium versus that of a socially-optimal solution. The idea of using worst-case analysis to study situations under competition was introduced by Koutsoupias and Papadimitriou (1999) and has gained followers over the last decade. The use of the price of anarchy as a metric of the welfare loss has been widely applied in economics to problems such as the study of competition and efficiency in congested markets (Acemoglu and Ozdaglar, 2007), games with serial, average and incremental cost sharing (Moulin, 2008), price and capacity competition (Acemoglu et al., 2009), and Vickrey-Clarke-Groves mechanisms (Moulin, 2009), resource allocation problems (Kelly, 1997; Johari and Tsitsiklis, 2004), and congestion games (Roughgarden and Tardos, 2004; Correa et al., 2008). In our setting, we define the welfare loss as the ratio between the social welfare of a socially-optimal solution—the sum of the principal's and agent's payoffs when the first-best effort level is chosen—and that of the subgame perfect equilibrium in which the principal offers the agent a performance-pay contract and then the agent chooses the effort level.

The goal of this paper is two-fold. First, we provide simple parametric bounds on the welfare loss of a given instance in the presence of limited-liability and moral hazard. These bounds allow one to directly quantify the inefficiency of a given instance without the need to determine its first-best and second-best effort levels, and additionally, they shed some light on the structure of instances with high welfare loss. Second, we study the worst-case welfare loss (i.e., the price of anarchy) among all instances with a fixed number of effort and outcome levels. The worst ratio is with respect to the parameters that define an instance of the problem: the outcome vector, the vector of agent’s costs of effort, and the probability distribution of outcomes for each effort level.

In order to obtain our bounds for the welfare loss, we assume throughout that the probability distribution of output, which is parameterized by the effort level, satisfies the monotone likelihood ratio property (MLRP)\(^3\)—under this property, a higher output is a better signal that the agent has chosen a higher effort level—and that the ratio of marginal cost to marginal probability of the highest outcome is non-decreasing with the effort level (IMCP).\(^4\) The latter assumption ensures that local incentive compatibility constraints are sufficient to induce an effort level, and is weaker than the well-known convexity of the distribution func-

\(^3\)The MLRP assumption is pervasive in the principal-agent literature (Grossman and Hart, 1983; Rogerson, 1985).

\(^4\)IMCP stands for Increasing Marginal Cost to marginal Probability. We thank an anonymous referee for suggesting the simplification.
tion (CDFC) that is also related to the idea of decreasing marginal returns to effort (see, for instance, Rogerson (1985) and Mirrlees (1999)). Furthermore, it is assumed that the sequence of prevailing social welfare under increasing effort levels is quasi-concave (QCSW); i.e., social welfare is single-peaked in the effort level.

Under our assumptions, inefficiency is introduced when the socially-optimal effort level is high but the optimal choice for the principal is to induce a low effort level. The principal may find it optimal to induce a low effort level because the gain of increasing the expected output is smaller than the cost of inducing the high effort level. Inducing a high effort level is costly for the principal, in turn, because of the agent’s limited liability constraint: the principal cannot severely punish the agent when a there is a realization of a bad outcome. Although he rewards the agent when there is a realization of a good outcome, that is not enough to compensate the cost incurred in ‘bad times.’

An unbounded welfare loss would arise if the social welfare of the high effort level were arbitrarily larger than that of the low effort level and, regardless, the principal prefers to induce the lower effort level. Because social welfare is the sum of the principal’s and agent’s utility, the latter implies that the agent is capturing most of the social welfare at the high effort level. Our results preclude an arbitrarily large welfare loss because, for any effort level, the limited liability rent given up cannot be arbitrarily larger than the principal’s own utility. Under assumptions MLRP, IMCP and QCSW, we establish that for any instance of the problem, the welfare loss is bounded above by a simple formula involving the probabilities of the highest possible outcomes (or alternatively the agent’s cost). The results arise from the fact that MLRP implies that the principal pays a bonus only when the highest outcome is observed.\(^5\) Our results suggest that the potential consequence of dealing with a moral-hazard problem only by using a performance-pay contract may have a non-negligible impact in the welfare of the system.

Subsequently, we show that the worst-case welfare loss among all instances with a fixed number of effort and outcome levels that satisfy our assumptions is equal to the number of effort levels \(E\). As a consequence, the social welfare of a subgame perfect equilibrium is guaranteed to be at least \(1/E\) of that of the social optimum. We prove that the worst-case is attained by a family of instances in which the likelihood of the highest outcome increases at a geometric rate with the effort level. Our results suggest that moral hazard is more problematic in situations where the agent’s available actions are more numerous and when the informational problem is such that the likelihood ratio of the highest outcome between the highest and the lowest effort level is higher. In a situation with high welfare loss, the principal—who realizes that a fraction of that loss could increase his rent—may engage in

\(^5\)In addition to this, Balmaceda et al. (2012) show that our results are robust by relaxing some assumptions of the model presented here.
monitoring with the purpose of limiting the agent’s action space, as well acquiring better
time information about the agent’s effort level. On the one hand, monitoring provides more
information and may decrease the agent’s rent needed to induce him to choose the desired
effort level. On the other hand, monitoring is costly. Thus, our results suggest that the
principal should engage in monitoring when the likelihood ratio mentioned above is large,
otherwise it should be optimal to pay a bonus only when the highest outcome is observed
and to punish the agent as much as possible when any other outcome is realized.

As an extension to the basic model, we study the welfare loss when contracts are restricted
to be linear. This is motivated by the prevalence of linear contracts in real life. In fact,
Salanié (2003, p. 474) concludes that “[t]he recent literature provides very strong evidence
that contractual forms have large effects on behavior. As the notion that ‘incentive matters’
is one of the central tenets of economists of every persuasion, this should be comforting to
the community. On the other hand, it raises an old puzzle: if contractual form matters so
much, why do we observe such a prevalence of fairly simple contracts?” Surprisingly, we
show that a similar bound on the welfare loss holds when the principal is restricted to choose
linear contracts, and that the worst-case welfare loss is again equal to the number of effort
levels $E$. Our results provide bounds on the welfare loss and do not shed light on whether,
for the same instance, the restriction to linear contracts increases or decreases welfare loss.

In another extension, we study the welfare loss when there are multiple identical and
independent tasks, and for each task the agent chooses between two effort levels. We give a
simple bound on the welfare loss involving the probability that all tasks are successful when
the agent exerts both a high and low effort level, and find that worst-case welfare loss is 2,
regardless of how many tasks the agent has to work on.

In related work, Demougin and Fluet (1998) derive the optimality of the bonus contract
in the setting considered here; i.e., with a discrete outcome space. Kim (1997) presents con-
ditions that guarantee the existence of a bonus contract that achieves a first-best allocation
under limited liability but for a continuous outcome and effort space. Our results extend this
work by quantifying the impact on efficiency when the first-best allocation is not achieved.
In this context, our work should be viewed as a preliminary step in a broader agenda of
how to quantify the welfare loss of moral hazard in different settings. Along those lines,
Babaioff et al. (2009, 2012) study a principal-agent problem with an approach similar to
ours. They introduce a combinatorial agency problem with multiple agents performing two-
effort-two-outcome tasks, and study the combinatorial structure of dependencies between
agents’ actions and the worst-case welfare loss for a number of different classes of action
dependencies. They show that this loss may be unbounded for technologies that exhibit
complementarities between agents, while it can be bounded by a small constant for technolo-
gies that exhibit substitutabilities between agents. Instead, our model deals with a single agent and its complexity lies in handling more sophisticated tasks, rather than the interaction between tasks.

The rest of the paper is organized as follows. In Section 2, we introduce the model with its main assumptions and present some preliminary results that will prove useful in the rest of the paper. Section 3 presents our main results on bounds for the welfare loss. Section 4 extends our results in several directions while Section 5 concludes with some remarks and future directions of study. Proofs can be found in the appendix.

2 The Principal-Agent Model

2.1 The Basic Setup

We consider a risk-neutral principal and a risk-neutral agent in a setting with $E \geq 2$ effort levels and $S \geq 2$ outcomes. The agent chooses an effort level $e \in E \triangleq \{1, \ldots, E\}$, incurring a personal nonnegative cost of $c_e$. Effort levels are sorted in increasing order with respect to costs; that is, $c_e < c_f$ if $e < f$. Thus, a higher effort level demands more work from the agent. We denote by $c = (c_1, \ldots, c_E)$ the vector of agent’s costs. The task’s output depends on a random state of nature $s \in S \triangleq \{1, \ldots, S\}$ whose distribution in turn depends on the effort level chosen by the agent. Each state has an associated nonnegative dollar amount that represents the principal’s revenue. We denote the vector of outputs indexed by state by $y = (y^1, \ldots, y^S)$. Without loss of generality, the outputs are sorted in increasing order: $y^s < y^t$ if $s < t$; hence, the principal’s revenues are higher under states with a larger index. Finally, we let $\pi^e_s$ be the common-knowledge probability of state $s \in S$ when the agent exerts effort level $e \in E$. The probability mass function of the outcome under effort level $e$ is given by $\pi^e_e = (\pi^e_1, \ldots, \pi^e_S)$. An instance of the principal-agent problem is characterized by the tuple $I = (\pi, y, c)$, and we denote by $I^{E,S} \subset \mathbb{R}^{E \times S} \times \mathbb{R}^S \times \mathbb{R}^E$ the set of all valid instances with $E$ effort levels and $S$ outcomes, that is, $\pi = \{\pi^e_e\}_{e=1}^E$ are probability vectors, and the outputs $y$ and costs $c$ are nonnegative and increasing.

Because the agent’s chosen effort level $e$ cannot be observed by the principal, he cannot write a wage contract based on it. However, the principal can write a contract that conditions payments on the output. The timing is as follows. First, the principal makes a take-it-or-leave-it offer to the agent that specifies a state-dependent wage schedule $w = (w^1, \ldots, w^S)$. The contract is subject to a limited liability (LL) constraint specifying that the wage must be nonnegative in every possible state. The LL constraint excludes contracts in which the

---

6Balmaceda et al. (2012) consider other assumptions, including continuous effort levels and some other relaxations.
agent ends up paying back to the principal; thus preventing the agent from covering losses with his own wealth. Second, the agent decides whether to accept or reject the offer, and if accepted, then he chooses the effort level $e \in \mathcal{E}$ before learning the realized state and incurs a personal cost $c_e$. Third, the random state $s \in \mathcal{S}$ is realized, the agent is paid the wage $w^s$ and the principal collects the revenue $y^s$. The agent should accept the contract if it satisfies an individual rationality (IR) constraint specifying that the contract must yield an expected utility to the agent greater than or equal to that of choosing the outside option, which is normalized to zero. Furthermore, he will choose the effort level $e \in \mathcal{E}$ by maximizing his expected payoff $\pi_e w - c_e$; that is, the difference between the expected wage and the cost of the effort level chosen.

The principal’s problem consists on choosing a wage schedule $w$ and an effort level $e$ for the agent that solve the following problem:

$$
\begin{align*}
    u^P & \triangleq \max_{e \in \mathcal{E}, w} \pi_e (y - w) \\
    \text{s.t. } & \pi_e w - c_e \geq 0, \quad \text{(IR)} \\
    & e \in \arg \max_{f \in \mathcal{E}} \{\pi_f w - c_f\}, \quad \text{(IC)} \\
    & w \geq 0. \quad \text{(LL)}
\end{align*}
$$

The objective function measures the difference between the principal’s expected revenue and payment, hence computing his expected profit. Constraints (IR) and (LL) were described earlier. The incentive compatibility (IC) constraints guarantee that the agent chooses the effort level that maximizes his own profit. As is standard in the literature, we assume throughout the paper that when the agent is indifferent between two or more effort levels, he always picks the one preferred by the principal. Results can be extended to hold under strong IC without further gain in intuition and much more cumbersome mathematical derivations.

Following Grossman and Hart (1983), one can equivalently formulate the principal’s problem as

$$
    u^P = \max_{e \in \mathcal{E}} \{\pi_e y - z_e\},
$$

where $z_e$ is the minimum expected payment incurred by the principal that induces the agent to exert effort level $e$. We denote by $u_e^P \triangleq \pi_e y - z_e$ the principal’s maximum expected utility when effort level $e$ is implemented, and by $\mathcal{E}^P$ the set of optimal effort levels for the principal. Hence, $u^P = \max_{e \in \mathcal{E}} \{u_e^P\}$ and $\mathcal{E}^P = \arg \max_{e \in \mathcal{E}} \{u_e^P\}$.

Exploiting that the effort level set is finite, we write the IC constraint explicitly to obtain the minimum payment linear program corresponding to $e$, which is independent of the
output $y$:

\[
\begin{align*}
(MPLP_e) \quad z_e &= \min_{w \in \mathbb{R}^S} \pi_e w \\
\text{s.t.} \quad \pi_e w - c_e &\geq 0, \quad \text{(IR)} \\
\pi_e w - c_e &\geq \pi_f w - c_f \quad \forall f \in \mathcal{E} \setminus e, \quad \text{(IC)} \\
w &\geq 0. \quad \text{(LL)}
\end{align*}
\]

We say that the principal implements effort level $e \in \mathcal{E}$ when the wage schedule $w$ is consistent with the agent choosing effort level $e$. For a fixed effort level $e$, constraints (IR), (IC), and (LL) characterize the set of feasible wages that implement $e$. The principal will choose a wage schedule belonging to that set that achieves $z_e$ by minimizing the expected payment $\pi_e w$. We denote an effort level as implementable if is attainable under some wage schedule, that is, its set of feasible wages is nonempty.

From the perspective of the system, the social welfare when implementing effort level $e \in \mathcal{E}$ is given $u_{SW}^e \triangleq \pi_e y - c_e$, which corresponds to the sum of the principal’s utility $\pi_e y - z_e$ and the agent’s utility $z_e - c_e$. Notice that since $z_e$ is the optimal objective value of $MPLP_e$, we have by the agent’s IR condition that $z_e \geq c_e$. Thus, the agent’s utility is always non-negative, and the social welfare is at least the principal’s utility; i.e., $u_{SW}^e \geq u_e^p$ for all effort levels $e \in \mathcal{E}$.

### 2.2 Assumptions

We assume that the probability distributions $\pi_e$ satisfy the well-known monotone likelihood-ratio property.\(^7\)

**Assumption 1. [MLRP]** The distributions $\{\pi_e\}_{e \in \mathcal{E}}$ verify that $\frac{\pi_s^e}{\pi_t^e} > \frac{\pi_s^f}{\pi_t^f}$ for all states $s > t$ and effort levels $e > f$.

The assumption MLRP is pervasive in the economics of information literature, and in particular in the principal-agent literature (see, e.g., Grossman and Hart (1983); Rogerson (1985)). This property ensures that the higher the observed level of output, the more likely it is the agent exerted a higher effort level.

Any distribution satisfying MLRP also satisfies the weaker first-order stochastic dominance (FOSD) property. Let $F_e^S \triangleq \sum_{s'=1}^S \pi_{s'}^e$ be the cumulative distribution function for effort level $e$. Rothschild and Stiglitz (1970) proved that for a fixed outcome the cumulative

\(^7\)Our results also hold under the weaker assumption that the largest likelihood ratio is verified by the highest outcome, that is, $\pi_s^e/\pi_t^e > \pi_s^f/\pi_t^f$ for all states $s < S$ and effort levels $e > f$. We thank an anonymous referee for suggesting this point.
distribution function is non-increasing in the effort level; or equivalently, \( F_f^s \geq F_e^s \) for all states \( s \) and effort levels \( e > f \). A simple consequence of this that plays an important role in our derivations is that probabilities for the highest outcome \( S \) are sorted in increasing order with respect to effort levels; i.e., \( \pi_f^S < \pi_e^S \) for \( e > f \). Note that in the case of two outcomes, MLRP and FOSD are equivalent.

The following ratios are central in the analysis that follows so we refer to them explicitly. We let the ratio of marginal cost to marginal probability of the highest outcome be equal to

\[
m_e \triangleq \begin{cases} 
\frac{c_e}{\pi_e^S} & \text{if } e = 1, \\
\frac{c_e - c_{e-1}}{\pi_e^S - \pi_{e-1}^S} & \text{if } e \geq 2.
\end{cases}
\]

In the case of two outcomes and an arbitrary number of effort levels, our results hold under the additional natural assumption that all effort levels are implementable. For the general case of an arbitrary number of outcomes we impose the following two additional assumptions.

**Assumption 2. [IMCP]** The ratio of marginal cost to marginal probability of the highest outcome is non-decreasing with the effort level; i.e., \( m_e \leq m_{e+1} \) for all effort levels \( 1 \leq e < E \).

This assumption establishes that the marginal cost of increasing the probability of outcome \( S \) from \( \pi_{e-1}^S \) to \( \pi_e^S \), expressed in terms of per-unit increased in probability, is non-decreasing with the effort level. This assumption is satisfied, for instance, when the marginal cost of effort is constant and there are decreasing marginal returns to effort. It is also satisfied when the marginal cost of effort is increasing and there are constant marginal return to effort. Hence, this assumption is the natural analogue of the standard convexity of the distribution function assumption (CDFC).\(^8\) In fact, one can easily show that CDFC together with convexity of the cost function implies IMCP. Our condition is weaker than CDFC since we only impose a restriction on the likelihoods for the highest outcome while the latter requires convexity of the distribution for every outcome.

The assumptions CDFC together with MLRP are standard conditions for the validity of the first-order approach (see, for instance, Rogerson (1985), Mirrlees (1974, 1999, 1976), Laffont and Martimort (2001) and Salanié (2003)).\(^9\) This is meant to ensure that critical points are global maxima, making local conditions enough to characterize the global optima for the agent’s effort level. Here, assumption IMCP plays a similar role. That is, an effort

\(^8\)CDFC requires that the distribution function is convex with respect to the effort level for any given fixed outcome; that is, \( F_{e+1}^s - F_e^s \geq F_e^s - F_{e-1}^s \) for all outcomes \( s \) and effort levels \( 1 < e < E \).

\(^9\)See, Jewitt (1988) for a less stringent but much less used condition ensuring that the first-order approach is valid. One advantage of Jewitt’s conditions is that they are closer to the economic notion of decreasing returns.
level satisfying the local incentive compatibility constraint is a global maximum in the agent’s optimization problem. This allows to order the effort levels in a manner that facilitates the computations upon deviations.\footnote{\textit{However as Rogerson (1985, p. 1362) points out, “if output is determined by a stochastic production function with diminishing returns to scale in each state of nature, the implied distribution function over output will not, in general, exhibit the CDFC.” Thus, the CDFC in general requires more than diminishing returns (see Conlon (2009) for a more detailed discussion).}}

**Assumption 3.** [QCSW] The sequence of prevailing social welfare under increasing effort levels is quasi-concave.

This assumption establishes that the social welfare is unimodal, or equivalently for some effort level $e^{SO}$, social welfare is monotonically non-decreasing for effort levels $e \leq e^{SO}$ and monotonically non-increasing for effort levels $e \geq e^{SO}$. This technical assumption excludes the possibility of local maxima in the social welfare and greatly simplifies our results.

### 2.3 Preliminaries

A critical step needed to derive our results involves characterizing the minimum expected payment $z_e$ incurred by the principal when inducing an effort level $e \in \mathcal{E}$. From the theory of linear programming, we know that the optimal wage schedule is verified by an extreme point of the constraint set. In general, the IC constraints of many other effort levels $f \in \mathcal{E} \setminus e$ may be binding at the optimal wage schedule; thus, not allowing to compactly characterize the minimum expected payment. While not needed,\footnote{\textit{For details, see Balmaceda et al. (2012).}} assumption IMCP greatly simplifies the analysis since it implies that the IC constraint of the immediately lower effort level is necessary and sufficient to implement effort level $e$. That is, to implement the effort level $e > 1$, the principal only needs to check that the agent’s utility for the effort level at hand dominates that of the previous effort level $e - 1$. Formal proofs of all the forthcoming results are placed in the appendix.

**Proposition 2.1.** Assume that MLRP and IMCP are verified. For every effort level $e \in \mathcal{E}$ the following holds.

\begin{enumerate}
  \item The minimum expected payment that the principal makes to the agent is $z_e = \pi^S_e m_e$.
  \item The optimal contract pays $w^s_e = 0$ for all $s < S$ and $w^S_e = m_e$ for $s = S$.
\end{enumerate}

The optimal wage schedule in Proposition 2.1 rewards the agent only when the highest outcome is realized. This result implies that for any implementable effort level, the optimal contract is a bonus contract of the pass/fail type that pays a bonus only when the outcome with the highest likelihood ratio is realized and pays the lowest possible payment allowed by...
the limited-liability constraint in all other states. Specifically, the principal will only reward the highest outcome (due to the MLRP), and he will impose the maximum penalty for all other outcomes. Because of the agent’s liability limit, this maximum penalty is equal to zero.

It is worthwhile to say a few words on the structure of the optimal contract. While the optimal payment scheme is very intuitive, from an empirical point of view it may be perceived as strange to pay a bonus only when the highest outcome is observed. The resulting threshold contract hinges upon (i) the monotone likelihood property, and (ii) the discreteness of the signal space. Indeed, this structure is well documented in the literature: Demougin and Fluet (1998) were the first to derive the optimality of bonus contracts in a discrete outcome space. In the case of a continuous outcome space, Kim (1997) showed that the optimal contract pays a fixed wage plus a bonus when the outcome exceeds a given threshold. He can prove this because he assumes that the first-best effort is achievable; otherwise the optimal contract would consist of an infinite payment in the zero-probability event that the highest outcome occurs. The literature on financial contracts with moral hazard and limited liability avoids the unappealing feature of the optimal contract derived here by considering monotonicity constraints, in which case a threshold contract (i.e., debt contract) would be recovered (Innes, 1990; Matthews, 2001; Poblete and Spulber, 2012). However, within the moral hazard with limited liability literature that deals with employment contracts, monotonicity constraints are usually disregarded and many papers focus on the two outcome case, in which case the bonus-type contract derived here is the natural solution. Hence, our paper deals with bounds to the welfare loss that relate to this latter literature. In future work, we will consider adding monotonicity constraints to our model to connect our insights to the broader literature.

To close this section with preliminary results, we show that a simple consequence of the previous result is that the agent’s utility is non-decreasing with respect to the effort level.

**Corollary 2.2.** Assume that MLRP and IMCP hold. The agent’s utility $z_e - c_e$ is non-decreasing in the effort level implemented by the principal; that is, $z_e - c_e \leq z_{e+1} - c_{e+1}$ for all $1 \leq e < E$.

This implies that to induce higher effort levels the principal must increase the agent revenue at a larger rate than the increase in cost experienced by the agent. As we will discuss in more detail later, this creates the incentive to the principal to induce a smaller-than-optimal effort level.
3 Bounding the Welfare Loss

3.1 The Welfare Loss

The goal of a social planner is to choose the effort level $e$ that maximizes the social welfare $u_{e}^{SW}$, which we have defined in Section 2.1 as the difference between the principal’s expected revenue and the agent’s cost of effort. Since preferences are quasi-linear, wages are a pure transfer of wealth between the principal and the agent, and the social planner is not concerned about them. The optimal social welfare is given by

$$u_{SO} \triangleq \max_{e \in \mathcal{E}} \{u_{e}^{SW}\},$$

and the set of first-best efficient effort levels is given by $\mathcal{E}_{SO} \triangleq \arg \max_{e \in \mathcal{E}} \{u_{e}^{SW}\}$.

For a given instance of the problem $I = (\pi, y, c) \in \mathbb{I}_{E,S}$ we quantify the welfare loss, denoted by $\rho(I)$, as the ratio of the social welfare under the socially-optimal effort level to that of the socially worst second-best effort level; that is,

$$\rho(I) \triangleq \frac{u_{SO}}{\min_{e \in \mathcal{E}^P} u_{e}^{SW}}. \quad (3)$$

Note that the minimum in the denominator corresponds to taking a worst-case perspective that reflects that among the possibly multiple outcomes one cannot know which one will materialize. The welfare loss is clearly at least one because the social welfare of any sub-game perfect equilibrium cannot be larger than the socially-optimal one.

The goal of this section is two-fold. First we aim to provide simple parametric bounds on the welfare loss of a given instance in the presence of limited-liability and moral hazard under the previous assumptions. Second we study the worst-case welfare loss among all instances with a fixed number of effort and outcome levels satisfying the previous assumptions, which is commonly referred to as the Price of Anarchy in the computer science literature (Nisan et al., 2007). The worst-case welfare loss of a class of instances $\mathcal{I} \subseteq \mathbb{I}_{E,S}$ is defined as the supremum of the welfare loss over all instances in the class: \footnote{While the worst-case inefficiency corresponding to a family of instances $\rho(I)$ has been dubbed the ‘price of anarchy,’ the computer science literature refers to the welfare loss corresponding to an instance $\rho(I)$ shown in (3) by the ‘coordination ratio.’ Further, note that the price of anarchy for a maximization problem such as the one we work with in this article is often defined as the inverse of the ratio in (4). We do it in this way so ratios and welfare losses point in the same direction. Other definitions such as the relative gap between the first-best and second-best solutions are possible, but we keep the standard one for consistency.}

$$\rho(\mathcal{I}) \triangleq \sup_{I \in \mathcal{I}} \rho(I). \quad (4)$$

In particular, we are interested in the worst-case welfare loss of the class of instances $\mathbb{I}_{E,S}^{E_{MIQ}} \subset \mathbb{I}_{E,S}$. 

I with E effort levels and S outcome levels satisfying assumptions MLRP, IMCP and QCSW.

It is worth noting that without limited liability, moral hazard does not affect the efficiency of the optimal contract. Namely, if the principal and the agent are risk-neutral and there is no limited liability constraint, the minimum expected payment \( z_e \) incurred by the principal when inducing a implementable effort level \( e \) is \( c_e \) (see, e.g., Laffont and Martimort (2001, p. 154)). Thus, without the constraint (LL), the principal implements the socially-optimal effort level and he fully captures all social surplus, leaving no rent to the agent. As a consequence, the worst-case welfare loss is 1.

### 3.2 The Main Result

Situations with high inefficiency may arise when the socially-optimal effort level is high, but the optimal choice for the principal is to induce a low effort level. In view of Corollary 2.2, inducing a high effort level is costly for the principal because of the agent’s limited liability rent, which must increase at a higher rate than the agent cost. As a result the principal may find it optimal to induce a low effort level because the gain of increasing the expected output is smaller than the cost of inducing the high effort level.

We bound the welfare loss when the effort level is chosen from a discrete set that takes \( E \geq 2 \) possible values, and give a bound on the welfare loss that involves the marginal return to effort given by

\[
\text{MRE}_e \triangleq \frac{\pi_{e}^{S} - \pi_{e-1}^{S}}{\pi_{e}^{S}},
\]

for effort levels \( 1 < e \leq E \). A salient feature of this bound, called MRE from now on, is that, in some instances, it may achieve a well-behaved limit as we increase the number of effort levels. Additionally, it provides some structural insights on instances with high welfare loss.

**Theorem 3.1.** Let \( I = (\pi, y, c) \in \mathbb{I}^{E, S}_{\text{MIQ}} \) be an instance satisfying assumptions MLRP, IMCP and QCSW. Then, the welfare loss

\[
\rho(I) \leq 1 + \sum_{e=2}^{E} \frac{\pi_{e}^{S} - \pi_{e-1}^{S}}{\pi_{e}^{S}}.
\]

The MRE bound can be alternatively written as \( E - \sum_{e=2}^{E} \pi_{e-1}^{S} / \pi_{e}^{S} \), which clearly implies that \( \rho(I) \leq E \). In words, the social welfare of a subgame perfect equilibrium is guaranteed to be at least \( 1/E \) of that of the social optimum. In Section 3.4 we offer an instance satisfying all the assumptions whose welfare loss is arbitrarily close to this upper bound, implying that this upper bound is tight (i.e., \( \rho(I^{E, S}_{\text{MIQ}}) = E \)). The insight of this result is that the welfare loss cannot be arbitrarily large because, for a given effort level, the limited-liability rent given up
cannot be arbitrarily larger than the principal’s own utility. To see this, let us consider the case of two effort levels. An unbounded welfare loss would imply that the social welfare under the high effort level is arbitrarily larger than that under the low effort level. To be in an inefficient situation the principal needs to prefer to induce the low effort level, which implies that his utility from implementing the high effort level is even smaller. Continuing with this line of reasoning, because social welfare is the sum of the principal and agent utilities, the agent utility for the high effort level has to be large since he is capturing most of the social welfare. But this cannot happen. Indeed, when proving our result we derive the inequality $z_e - c_e \leq u_e^{SW} - (u_e^{SW} - u_{e-1}^{SW})/\text{MRE}_e$, which implies the higher the difference in social welfare between the high and low effort levels, the lower the agent utility at the high effort level. As a result the agent utility at the high effort level is bounded, yielding a contradiction to the argument above.

Let $e^{SO} = \max E^{SO}$ be the greatest effort level that maximizes social welfare and $\bar{e} \triangleq \max\{e \in E : z_e = c_e\}$ the greatest effort level that leaves no rent to the agent. When $e^{SO} < \bar{e}$, the welfare loss is one since the principal has no incentive to implement any effort level higher than the optimal one because the agent’s utility is non-decreasing with the effort level by Corollary 2.2. In contrast when the opposite occurs, the fact that the principal induces the agent to implement the preferred effort level instead of the socially-optimal one is costly for the system. The condition $e^{SO} < \bar{e}$ is the discrete analogue of the one given by Proposition 1 in Kim (1997) guaranteeing the existence of a bonus contract that achieves a first-best allocation in the case of continuous effort levels and outcomes.

After some algebra (see the proof of Theorem 3.1), the MRE bound itself can be upper bounded by the following compact expression

$$1 + \ln(\pi_e^{S}/\pi_1^S)$$

that just depends on the probabilities of the lowest and highest effort levels but not on the number of effort levels. This is appealing because one has an upper bound on the inefficiency that only depends on the basic parameters of the instance at hand, without having to compute first or second-best contracts.

Exploiting our characterization of the minimum expected payments, it is straightforward to see that one can also upper bound the MRE bound by\(^\text{13}\)

$$1 + \sum_{e=2}^{E} \frac{c_e - c_{e-1}}{c_e} < 1 + \ln \left( \frac{c_E}{c_1} \right).$$

\(^{13}\)This result follows from the fact that for every effort level $e > e^* \geq 1$ we have by Proposition 2.1 that $\pi_e^S m_e \geq c_e$, or equivalently $(c_e - c_{e-1})/c_e \geq (\pi_e^S - \pi_{e-1}^S)/\pi_e^S$.\text{\Large 14}
The first bound in terms of the marginal cost of effort is tight (see, e.g., the instance in Section 3.4). This bound was first derived by Bastin et al. (2013), extending results of Balmaceda et al. (2012).

3.3 A Welfare Bound for 2 Outcomes and $E$ Effort Levels

In the case of two outcomes and an arbitrary number of effort levels, we can replace the earlier assumptions by weaker ones. For the MRE bound to hold, we only require that all effort levels are implementable together with FOSD.

**Assumption 4. [IMP]** All effort levels $e \in \mathcal{E}$ are implementable by the principal.

In the case of two outcomes, the assumption IMP guarantees that the ratios of marginal cost to marginal probability satisfy that $m_e \leq m_{e+1}$ for all effort levels $1 < e < E$, which is almost equivalent to IMCP with the exception of $e = 1$. Under two outcomes, this can be shown to imply that the social-welfare is quasi-concave. Since IMCP may not hold, it may be the case that for some effort level $e > 1$ the IR constraint is binding instead of the LL constraint, and the principal gives up no limited-liability rent. Nevertheless, we can work around this issue and extend the characterization of the minimum expected payments to reflect this fact.

With this in mind, we let $\mathcal{I}_{E,2}^{F,I} \subset \mathcal{I}_{E,2}^{I}$ be the class of instances with $E$ effort levels and 2 outcome levels satisfying the weaker assumptions FOSD and IMP. The following theorem together with the instance provided in Section 3.4 yield that the worst-case welfare loss of this class is $\rho(\mathcal{I}_{E,2}^{F,I}) = E$. Notice that although $\mathcal{I}_{E,2}^{F,I} \supset \mathcal{I}_{M,I,Q}^{E,2}$, the worst-case welfare losses coincide in both cases, meaning that relaxing those assumptions does not degrade the worst-case.

**Theorem 3.2.** Let $I = (\pi, y, c) \in \mathcal{I}_{E,2}^{F,I}$ be an instance with two possible outcomes satisfying assumptions FOSD and IMP. Then, the welfare loss $\rho(I) \leq 1 + \sum_{e=2}^{E} (\pi_{e}^{S} - \pi_{e-1}^{S})/\pi_{e}^{S}$.

3.4 Tightness of the Bounds

In this section, we present instances satisfying the given assumptions whose welfare loss is arbitrarily close to the upper bound on the number of effort levels. More formally, we will define a family of instances $I^\varepsilon \in \mathcal{I}_{M,I,Q}^{E,2}$ with $E > 1$ effort levels and $S = 2$ outcomes parameterized by $\varepsilon \in (0, 1)$ satisfying assumptions MLRP, IMCP and QCSW that verify that $\lim_{\varepsilon \to 0} \rho(I^\varepsilon) = E$. Obviously, $I^\varepsilon \in \mathcal{I}_{E,2}^{F,I}$ as well.

The MRE bound on Theorem 3.1 provides some insight on the structure of the worst-case instance. For the bound to be close to the number of effort levels it should be the case that MRE should be as large as possible. Using the fact that the arithmetic mean dominates the
geometric mean (AM-GM inequality) and canceling terms we obtain that the MRE bound is upper bounded by

$$1 + \sum_{e=2}^{E} \frac{\pi^S_e - \pi^S_{e-1}}{\pi^S_e} = E - (E - 1) \left( \frac{1}{E - 1} \sum_{e=2}^{E} \frac{\pi^S_{e-1}}{\pi^S_e} \right) \leq E - (E - 1) \left( \frac{\pi^S_E}{\pi^S_1} \right)^{1 - \frac{1}{E - 1}}. $$

Because the equality between geometric and arithmetic means holds when all terms are constant, we obtain that the worst case likelihood of the highest outcome should increase at a geometric rate. Additionally, for the bound on the number of effort levels to be tight it needs to be the case that the likelihood ratio $\pi^S_E/\pi^S_1$ goes to infinity. In the remainder on this section we provide an instance exhibiting these properties.\(^{14}\)

First, fixing $0 < \varepsilon < 1$, we let the probabilities of the outcomes associated to each effort level be $\pi_e = (1 - \varepsilon^{E-e}, \varepsilon^{E-e})$ for $e \in \mathcal{E}$. Notice that the likelihood of the highest outcome increases at a geometric rate of $\varepsilon^{-1}$. The probability distributions are such that effort level $E$ guarantees a successful outcome with probability one, while the lower effort levels generate a failed outcome with high probability. Clearly, these distributions verify that $\pi_2 \leq \ldots \leq \pi_E$, and thus they satisfy MLRP. (Recall that in the case of two outcomes MLRP and FOSD are equivalent.)

For the agent costs we let $c_E = \varepsilon^{-E}$, and then indirectly define the costs by imposing that the agent’s utility is $z_e - c_e = e - 1$ for all $e \in \mathcal{E}$. Assuming that IMCP holds to derive the required conditions of the instance, our characterization of the minimum expected payments yields that $z_e = \pi^2_e m_e$, and we obtain $c_{e-1} = c_e \varepsilon - (e - 1) (1 - \varepsilon)$ for $e = 2, \ldots, E$. Notice that this implies that $m_{e+1} - m_e = 1/\varepsilon^{E-\epsilon}$ and the instance satisfies the IMCP assumption. Finally, let the output be $y = (0, E + \varepsilon^{-E})$. One can prove inductively that the social utility is $u^\text{SW}_e = e + \sum_{i=1}^{E-e} \varepsilon^i$, and that principal’s utility is $u^\text{P}_e = \sum_{i=0}^{E-e} \varepsilon^i$, for $e \in \mathcal{E}$. Hence, the instance fulfills assumption QCSW because $u^\text{SW}_1 \leq \ldots \leq u^\text{SW}_E$ and the principal’s utilities satisfy $u^\text{P}_1 \geq \ldots \geq u^\text{P}_E$, so it is optimal for the principal to implement effort level 1. Additionally, it is not hard to show that costs are increasing, non-negative, and that they diverge to infinity as $\varepsilon$ goes to zero.

The welfare loss corresponding to this instance is given by

$$\rho(I^\varepsilon) = \frac{u^\text{SW}_E}{u^\text{SW}_1} = \frac{E}{1 + \sum_{i=1}^{E-1} \varepsilon^i} = E \frac{1 - \varepsilon}{1 - \varepsilon^E},$$

which converges to $E$ as $\varepsilon \to 0^+$. Notice that the MRE bound of Theorem 3.1 gives $E - \varepsilon(E - 1)$ which goes to $E$ as $\varepsilon$ goes to zero, as expected. Therefore, the upper bound on the number of effort levels is tight because we found a series of instances converging to a

\(^{14}\)We thank an anonymous referee for suggesting this argument.
matching upper bound.

Finally, for a fixed number of effort levels the bound on the welfare-loss given by \(1 + \log\left(\frac{\pi_E}{\pi_1}\right)\) is not tight in general, except in the trivial case when the welfare loss is one. However, if we are allowed to increase the number of effort levels arbitrarily, while holding fixed the likelihood ratio \(r = \frac{\pi_E}{\pi_1}\), then it is possible to construct instances with likelihood ratio equal to \(r\) whose welfare loss converges to \(1 + \ln(r)\) as the number of effort levels grows to infinity.

## 4 Extensions

In this section we look at different extensions of our results which illustrate the flexibility of our approach. We establish that the worst-case welfare loss does not change if we restrict the contracts to be linear, which are prevalent in practice. Then, we show how our bounds can be tightened if the agent has to work on many identical tasks and for each one he has to choose one of two effort levels. Although, there are many choices for effort levels on the aggregate agency problem, we show that the worst-case welfare loss is two. Finally, in a companion working paper we establish that our results generalize to arbitrary (potentially negative) costs for effort, to the case when the outside option has nonzero utility, and to more general limited-liability constraints, for which we can provide a more refined parametric bound that captures the dependence between the inefficiency of a contract and the minimum wage (Balmaceda et al., 2012).

### 4.1 Linear Contracts

Linear contracts are simple to analyze and implement, are observed in many real-world settings, and have an appealing property, which is to create uniform incentives in the following sense. Think of \(y\) as aggregate output over a given period of time (say a year), and think of the agent taking several actions during this period (say one per day). In this setting, a non-linear contract may create unintended incentives over the course of the year, depending on how the agent has done so far. For instance, suppose the contract pays a bonus if the output exceeds a given target level. Given this contract, once the agent reaches the target, he will stop working. He will also stop working if he is far from reaching the target in a date close to the end of the year. Neither of these two things will occur when the agent faces a linear incentive contract.\(^\text{15}\)

\(^{15}\)A growing body of evidence is consistent with the prediction that non-linear contracts create history-dependent incentives: see Healy (1985) on bonus plans with ceilings and floors, Asch (1990) and Oyer (1998) on bonuses tied to quotas, and Brown et al. (1996) and Chevalier and Ellison (1999) on how the convex relationship between mutual fund performance and assets under management caused risk-taking portfolio choices by ostensibly conservative funds. For a nice theoretical discussion about this point see Holmström
However, focusing on linear contracts is not free of problems. Mirrlees (1999) showed that the best linear contract, \( w = a + by \), is worse than various non-linear contracts. Why, then, are linear contracts so common in practice? A principal could pay a large premium for “simplicity” if he adopts a linear contract in settings where nonlinear contracts are optimal. A partial explanation arises by comparing the welfare loss when the principal attempts to motivate a risk-neutral worker subject to limited liability with a linear contract rather than an unrestricted one. Naturally, this does not fully answer the question since this would necessitate a tight bound on the principal’s profit loss. A plausible explanation of why firms do not deviate from linear contracts may be the fact that non-linear contracts are expensive to manage, lend themselves to gaming, and the ensuing worst-case bound is the same as that when the optimal nonlinear contract is used.

A linear contract is characterized by two parameters, an intercept \( a \) and a slope \( b \), such that the wage at state \( s \in S \) is given by \( w_s = a + bs \). The parameters of the contracts are not restricted in any way other than by the limited liability constraint. Given an instance \( I \in \Pi^{E,S} \) we denote by \( \rho^L(I) \) the welfare loss when the principal implements the optimal linear contract. The next result bounds the welfare loss under linear contracts under assumptions FOSD and IMP\(_L\), where IMP\(_L\) imposes that all effort levels \( e \in E \) are implementable by the principal via a linear contract.

**Theorem 4.1.** Let \( I = (\pi, y, c) \in \Pi_{FIL}^{E,S} \) be an instance restricted to linear contracts satisfying assumptions FOSD and IMP\(_L\). Then, the welfare loss

\[
\rho^L(I) \leq 1 + \sum_{e=2}^{E} \frac{\pi_e y - \pi_{e-1} y}{\pi_e y - y}. 
\]

The proof proceeds by exploiting that an instance \( I \in \Pi^{E,S} \) restricted to linear contracts with an arbitrary number of outcomes and effort levels can be reduced to an unrestricted problem \( \tilde{I} \in \Pi^{E,2} \) with the same number of effort levels but only two outcomes (one for each parameter). As a result we obtain that the welfare loss restricted to linear contracts is equal to the unrestricted welfare loss of the reduced instance, or equivalently \( \rho^L(I) = \rho(\tilde{I}) \). In view of the result in Section 3.3 for instances with two outcomes, we only need IMP and FOSD for our results to hold, which we show they translate to the reduced instance. Alternatively, we could impose that IMCP holds for the reduced instance instead of IMP, however, we believe that our assumption on the implementability of the effort levels is more natural for this setting.

Next, we characterize the worst-case welfare loss restricted to linear contracts of the class of instances \( \Pi_{FIL}^{E,S} \subset \Pi^{E,S} \) with \( E \) effort levels and \( S \) outcome levels satisfying assumptions and Milgrom (1987).
FOSD and IMP\textsuperscript{L}. The previous theorem yields that $\rho^L(I_{FIL}^{E,S}) \leq E$. When the original problem has two outcomes there is a one-to-one correspondence between any wage schedule and the two parameters of the linear contract. Hence, the instance in Section 3.4 applies and the bound for welfare loss is tight. As a result we obtain that the worst-case welfare loss restricted to linear contracts is $\rho^L(I_{FIL}^{E,S}) = E$, as before.

Recall that in Section 3, we proved bounds for the welfare loss under non-linear contracts. Combining those bounds with Theorem 4.1, we have that the worst-case welfare loss under linear contracts coincides with that of the general case. Therefore, our results suggest that with respect to the worst-case scenario restricting the attention to linear contracts does not generate more inefficiency from a social point of view. Our results, however, do not shed light on whether, for a fixed instance $I$, the unrestricted welfare loss $\rho(I)$ and the one restricted to linear contracts $\rho^L(I)$ coincide or not.

To put our bounds on perspective we conclude by comparing the unrestricted MRE bound of Theorem 3.1 with that parametric bound restricted to linear contracts of Theorem 4.1. Let the marginal return to effort for linear contracts be defined as

$$MRE_e^L \triangleq \frac{\pi_e y - \pi_{e-1} y}{\pi_e y - y'},$$

for effort levels $1 < e \leq E$. If we strengthen FOSD to MLRP we can show that $MRE_e^L \leq MRE_e$ for effort levels $e > 1$, and thus the same MRE bound of Theorem 3.1 applies in the case of linear contracts.\textsuperscript{16}

### 4.2 Multiple Tasks

Often principal-agent relationships require that the agent performs different tasks, each endowed with different actions. In this section, we consider a principal-agent relationship with multiple tasks, adopting a model proposed by Laux (2001). The principal is endowed with $N$ identical and stochastically independent tasks. Each task has two possible outcomes: success or failure. The corresponding payoffs for the principal are $y$ if the task is successful, and $y'$ in the case of failure, with $y > y'$. For each task, the agent can exert two effort levels: either high or low. The high effort level entails a cost $c_h$ for the agent, while the cost of the low effort level is $c_l$. Since a higher effort level demands more work, $c_h > c_l$. Finally, we denote by $p_h$ the probability of success when the effort level is high, and by $p_l$ the probability of success when the effort level is low. We assume that FOSD holds for each task, from where $p_h > p_l$ (the higher the effort level, the greater the likelihood of success).

\textsuperscript{16}Letting $y' = y - y'1$, the condition $MRE_e^L \leq MRE_e$ can be written after some algebra as $(\pi_e^S \pi_{e-1} - \pi_{e-1}^S \pi_e)y' \geq 0$, which follows because $y' \geq 0$ from outputs being increasing, and $\pi_e^S \pi_{e-1} \geq \pi_{e-1}^S \pi_e$ from MLRP.
The principal hires an agent to perform the $N$ tasks. Since tasks are identical, the principal offers a compensation that depends only on the number of tasks that end up being successful, denoted by $s \in \mathcal{S} = \{0, \ldots, N\}$; the identity of each task is irrelevant. Hence, the agent is paid a wage $w^s$ when $s$ tasks turn out to be successful. The total revenue for the principal is thus $y^s = s\overline{y} + (N - s)y$ for $s \in \mathcal{S}$. In view of the tasks’ symmetric nature, the agent is indifferent between tasks and he is only concerned about the total number of tasks in which he exerts high effort level. We define the aggregated effort level $e \in \mathcal{E} = \{0, \ldots, N\}$ as the number of tasks in which the agent works hard. Notice that, for notational simplicity, we adopt indices that start at zero for both effort levels and states. We assume that costs are additive, and linear in the number of tasks. Hence, the aggregate costs for the agent are $c_e = e c_h + (N - e) c_l$ for $e \in \mathcal{E}$. Finally, note that the probability of having $s$ successful tasks, given that the agent works hard on $e$ tasks, is given by

$$
\pi^s_e = \sum_{i=0}^{s} \binom{e}{i} p_h^i (1 - p_h)^{e-i} \binom{N - e}{s - i} p_l^{s-i} (1 - p_l)^{N - e - s + i},
$$

where we let $\binom{n}{k} = 0$ if $k > n$ for notational simplicity.

This model can be fully reduced to a principal-agent model with a single task, $N + 1$ states and $N + 1$ effort levels. To map the multiple-task model into the model of Section 2 we show that the aggregate instance satisfies MLRP, which implies that the larger the observed number of successful tasks, the more likely it is that the agent works hard in many tasks.

**Lemma 4.2.** Assume that FOSD holds for each of the identical tasks. When the principal hires one agent to perform $N$ of these independent tasks, the probability distribution of the outcome satisfies MLRP for the aggregated problem.

Because effort levels are aggregated in $N + 1$ levels, our previous results would indicate that the welfare loss is upper bounded by $N + 1$. However, that bound is not tight. Laux (2001) shows that when the agent exerts high effort level in all tasks (aggregated effort level $N$), the only binding constraint is the one that corresponds to choosing a low effort level for the $N$ tasks (aggregate effort level $0$). As a result, in equilibrium the principal induces either the lowest or highest aggregate effort level, which in turn implies that $2$ is a tight bound for the welfare loss. Note that assumptions IMCP and QCSW are not needed for the following result to hold. The latter stems from the fact that social welfare is linear with the number of tasks for which the agent exerts a high effort level.

**Theorem 4.3.** Assume that FOSD holds for each task. In the principal-agent problem in which both players are risk-neutral and there are $N$ identical and independent tasks, welfare
loss is bounded by $2 - (p_l/p_h)^N < 2$.

5 Conclusions

This paper quantifies the welfare loss that arises from the principal’s inability to observe the agent’s effort level when there is limited liability. We have provided a simple parametric bound on the welfare loss involving the probabilities of the highest possible outcome. This bound leverages the structure of the optimal contract that pays a bonus only when the highest outcome is observed. The general structure of the bounds found in this paper suggests that the welfare loss in a principal-agent relationship depends on the set of effort levels available to an agent. In complex jobs where many actions are available to the agent, the welfare losses could be potentially high. This conclusion is valid under a variety of assumptions on the primitives of the model.

The principal-agent model in its different forms has been used to explain many contractual arrangements such as sharing contracts, insurance contracts, managerial contracts, political relationships, etc. In addition, it has been used to provide an economic theory of the firm and a theory of organizational forms. Our results show that in these cases and in many others, the existence of an agency relationship with moral hazard may have nontrivial consequences in terms of welfare loss and thus the proper design of contracts and organizations to deal with moral hazard is of great practical importance.

Nonetheless the relevance of our results, they open many more questions than they answer. The most ambitious question will be to quantify the welfare loss for each plausible instance with respect to the cost function, the monitoring technology and utility functions. We have only provided a bound for the worst-case welfare loss in the particular case in which neither the principal nor the agent is risk averse. Hence, this paper is just a preliminary approach to the hard question of how to quantify the welfare losses from moral hazard under different scenarios.

A next feasible step to keep progressing on the hard question of how to quantify the welfare loss is to consider a risk-averse agent. In this case the optimal contract is highly nonlinear, and thus its characterization in terms of the main parameters is a complex task. There is an exception to this, which is given by the linear agency model by Holmström and Milgrom (1987), which could used as the starting point to investigate the question at hand. Another potential extension involves considering monotonicity constraints as done in the literature on optimal financial contracts with moral hazard. Monotonicity of the contract and the principal’s payoff with the outcome leads to more reasonable threshold contracts, at the expense of a more complex analysis.
Acknowledgements

We greatly appreciate the comments of two anonymous referees that helped us significantly improve the presentation of the paper. This work was partially supported by FONDECYT Chile through grants 1140140 and 1130671, by the Instituto Milenio Sistemas Complejos de Ingeniería, by the Millenium Nucleus Information and Coordination in Networks ICM/FIC P10-024F, and by CONICET Argentina Grant PIP 112-201201-00450CO, Convenio Coopera-ción CONICET-CONICYT Resolución 1562/14, and FONCYT Argentina Grant PICT 2012-1324.

References


A Proofs

A.1 Proof of Proposition 2.1

Proof. The proof proceeds by considering the discrete counterpart of the relaxation approach of Rogerson (1985), which consists on relaxing the principal’s problem to include only the IC constraint for the immediately lower effort level \( e-1 \). Because we are expanding the constraint set, the objective value of the relaxed problem is less or equal than that of the original problem. Then we employ the Karush-Kuhn-Tucker conditions to characterize the optimal solution of the relaxed problem (which are sufficient because the problem is linear). Optimality for the original problem follows from proving that the optimal solution of the relaxed problem is feasible for the original problem, that is, it satisfies the IC constraints for every effort level. IMCP plays a critical role in this last part.

Fix an effort level \( e > 1 \). The relaxed problem is given by

\[
\begin{align*}
\min_{w \in \mathbb{R}^S} & \quad \pi_e w \\
\text{s.t.} & \quad \pi_e w - c_e \geq 0, \quad (5a) \\
& \quad \pi_e w - c_e \geq \pi_{e-1} w - c_{e-1}, \quad (5b) \\
& \quad w \geq 0. \quad (5c)
\end{align*}
\]

Introducing Lagrange multipliers \( \gamma \geq 0 \) for the IR constraint, \( \delta \geq 0 \) for the IC constraint, and \( \lambda \in \mathbb{R}^S \) with \( \lambda \geq 0 \) for the LL constraints we obtain the Lagrangian given by

\[
L_e = \pi_e w - \gamma (\pi_e w - c_e) - \delta (\pi_e w - c_e - \pi_{e-1} w + c_{e-1}) - \lambda w.
\]

The KKT conditions state that an optimal solution should satisfy the first-order conditions with respect to \( w^s \) as given by

\[
\frac{\partial L_e}{\partial w^s} = \pi^s_e - \gamma \pi^s_e - \delta (\pi^s_e - \pi^s_{e-1}) - \lambda^s = 0,
\]

(6)

together with constraints (5a), (5b) and (5c), and the complementarily slackness conditions.

We will show that \( w^* = m_e 1^S \) satisfies the optimality conditions of the relaxed problem. We begin with feasibility. For the IR constraint (5a) we obtain that \( \pi_e w^* - c_e = \pi^S_e m_e - c_e = \pi^S_e (m_e - m_{e,0}) \geq 0 \) by item (i) in Lemma A.1. For the IC constraint (5b) we obtain that \( (\pi_e - \pi_{e-1})w^* + c_{e-1} - c_e = (\pi^S_e - \pi^S_{e-1})m_e + c_{e-1} - c_e = 0 \) from our definition of \( m_e \). The LL constraint (5c) follows trivially.

For the complementarity slackness conditions to hold we need that: \( \gamma = 0 \) for the IR constraint, \( \delta \geq 0 \) because the IC constraint is binding, and the LL constraints imply that
\( \lambda^s \geq 0 \) for \( s < S \) and \( \lambda^S = 0 \). The first-order condition (6) for \( s = S \) gives that \( \delta = \pi^S_e / (\pi^S_e - \pi^S_{e-1}) \), which is clearly non-negative from FOSD. The same condition for \( s < S \) gives that \( \lambda^s = \pi^S_e - \delta(\pi^S_e - \pi^S_{e-1}) \). The non-negativity of the multiplier \( \lambda^s \) follows from MLRP between effort levels \( e, e-1 \) and outcome levels \( s, S \). Thus the proposed solution is optimal for the relaxed problem.

Feasibility of the relaxed solution to the original problem follows from checking that the IC constraints for effort levels \( f \neq e-1 \) hold. For \( f < e \) we obtain that \( (\pi_e - \pi_f)w^* + c_f - c_e = (\pi^S_e - \pi^S_f)m_e + c_f - c_e \geq (\pi^S_e - \pi^S_f)m_{e,f} + c_f - c_e = 0 \) where the inequality follows from by item (i) in Lemma A.1 and FOSD. The IC constraint for effort levels \( f > e \) holds similarly from item (ii) in Lemma A.1 and FOSD.

The result for effort level \( e = 1 \) follows similarly by considering a relaxed problem with no IC constraints and exploiting \( w^* = m_{11}^S \) as a candidate optimal solution.

### A.2 Additional Results

A critical consequence of IMCP is that only the lower adjacent incentive constraint matters. That is, when the principal wishes to implement effort level \( e \), he must be concerned only with a deviation towards effort level \( e-1 \). The following lemma is critical to show local incentive compatibility.

#### Lemma A.1

Assume that IMCP holds. Let

\[
m_{e,f} = \frac{c_e - c_f}{\pi^S_e - \pi^S_f}
\]

with the convention that \( c_0 = \pi^S_0 = 0 \) so that \( m_{e,0} = c_e/\pi^S_e \). We have that for all \( e \in E \):

(i) \( m_e \geq m_{e,f} \), for all \( f = 0, \ldots, e-1 \);

(ii) \( m_e \leq m_{f,e-1} \), for all \( f = e, \ldots, E \).

**Proof.** For item (i) note from IMCP we have that \( m_{g,g-1} \leq m_{e,e-1} \) for all \( g \leq e \), which implies that \((c_g - c_{g-1})(\pi^S_g - \pi^S_{g-1}) \leq (c_e - c_{e-1})(\pi^S_e - \pi^S_{e-1})\) because costs and the probabilities of the highest outcome are non-decreasing with respect to the effort level. Adding over \( g = f + 1, \ldots, e \) and collecting terms we obtain that

\[
(c_e - c_f)(\pi^S_e - \pi^S_{e-1}) \leq (c_e - c_{e-1})(\pi^S_e - \pi^S_{f}),
\]

and the result follows.

For item (ii) we proceed similarly and get from IMCP that \((c_g - c_{g-1})(\pi^S_g - \pi^S_{g-1}) \geq (c_e - c_{e-1})(\pi^S_g - \pi^S_{g-1})\) for \( g > e \). Adding over \( g = e + 1, \ldots, f \) and collecting terms we obtain
that
\[(c_f - c_e)(\pi^S_{e} - \pi^S_{e-1}) \geq (c_e - c_{e-1})(\pi^S_{f} - \pi^S_{e}),\]
and the result follows. \(\blacksquare\)

### A.3 Proof of Corollary 2.2

**Proof.** Observe that
\[z_e - c_e = \pi^S_e m_e - c_e \leq \pi^S_e m_{e+1,e} - c_e = \pi^S_{e+1} \frac{c_{e+1} - c_e}{\pi^S_{e+1} - \pi^S_e} + (c_{e+1} - c_e) - c_{e+1}\]
\[= \pi^S_{e+1} \frac{c_{e+1} - c_e}{\pi^S_{e+1} - \pi^S_e} - c_{e+1} = \pi^S_{e+1} m_{e+1} - c_{e+1} = z_{e+1} - c_{e+1},\]
where the inequality follows from property (ii) of Lemma A.1. \(\blacksquare\)

### A.4 Proof of Theorem 3.1

**Proof.** In order to prove the result we first show that the principal has no incentive to implement an effort level higher than \(e^{SO} = \max E^{SO}\) defined as the largest of the socially-optimal effort levels. Then, we obtain the bound by lower bounding the social welfare under the optimal effort level for the principal, which we denote by \(e^*\). We conclude by proving the logarithm bound for MRE.

**Step 1.** Because the utility of the agent increases with his effort level, the principal has no incentive to implement any effort level higher than \(e^{SO}\). Indeed, for \(f > e^{SO}\), Corollary 2.2 together with the fact that \(e^{SO}\) is the largest socially-optimal effort level imply that
\[u^P_{e^{SO}} - u^p_f = u^{SO} - u^{SW}_f + (z_f - c_f) - (z_{e^{SO}} - c_{e^{SO}}) > 0.\]
Hence, effort levels larger than \(e^{SO}\) provide a suboptimal utility to the principal, and can be disregarded.

**Step 2.** Note that if \(e^* = e^{SO}\) then the welfare loss is one and the bound trivially holds. In the remainder of the proof we shall consider the case where \(e^* < e^{SO}\). Consider an effort level \(e\) such that \(e^* < e \leq e^{SO}\). The fact that the social welfare is at least the principal’s utility, and Proposition 2.1 (because \(e > 1\)) imply that
\[u^{SW}_{e^*} \geq u^p_{e^*} \geq u^p_e = \pi^S_e y - z_e = \pi^S_e y - \pi^S_e \frac{c_e - c_{e-1}}{\pi^S_e - \pi^S_{e-1}}.\]
We now use that $c_e - c_{e-1} = \pi_e y - \pi_{e-1} y - u^{SW}_{e} + u^{SW}_{e-1}$ and obtain that the last expression is equivalent to
\[
\frac{\pi_{e-1} \pi_e - \pi_e \pi_{e-1}}{\pi_e - \pi_{e-1}} y + \frac{u^{SW}_{e} - u^{SW}_{e-1}}{\pi_e - \pi_{e-1}} \pi_e S \geq \frac{\pi_e S}{\pi_e - \pi_{e-1}} (u^{SW}_{e} - u^{SW}_{e-1}),
\] (8)
where the inequality follows from MLRP because $\pi_{e-1} \pi_e S \geq \pi_e \pi_{e-1} S$. The last inequality chain is equivalent to $(\pi_e S - \pi_{e-1} S)/\pi_e S \geq (u^{SW}_{e} - u^{SW}_{e-1})/u^{SW}_{e}$. Summing over the effort levels $e = e^* + 1, \ldots, e^{SO}$ and rearranging terms we get that
\[
\frac{u^{SW}_{e^{SO}}}{u^{SW}_{e^*}} \leq 1 + \sum_{e=e^*+1}^{e^{SO}} \frac{\pi_e S - \pi_{e-1} S}{\pi_e S} \leq 1 + \sum_{e=2}^{E} \frac{\pi_e S - \pi_{e-1} S}{\pi_e S},
\]
where the second inequality follows from including all effort levels greater than one in the sum, and using the fact that the probabilities $\pi_e S$ are increasing with the effort level.

**Step 3.** Finally, using the fact that $\ln(x) \leq x - 1$ for all $x \geq 0$, we can bound each term of the MRE bound by $-\ln \left( \frac{\pi_{e-1} S}{\pi_e S} \right)$, which yields
\[
1 + \sum_{e=2}^{E} \frac{\pi_e S - \pi_{e-1} S}{\pi_e S} \leq 1 - \sum_{e=2}^{E} \ln \left( \frac{\pi_{e-1} S}{\pi_e S} \right) = 1 + \ln \left( \frac{\pi_{e^{SO}} S}{\pi_{e^*} S} \right). \quad \square
\]

**A.5 Proof of Theorem 3.2**

**Proof.** We will prove the result in four steps. First, we show that, in the case of two outcomes, the fact that every effort level $e \in \mathcal{E}$ is implementable implies that $m_e \leq m_{e+1}$ for all effort levels $1 < e < E$ (note that this is IMCP restricted to effort levels greater than one). We denote this restricted condition as R-IMCP. Second, we prove that R-IMCP implies that social-welfare is quasi-concave and QCSW automatically holds. Third, we characterize the optimal payments made by the principal under R-IMCP. Under this condition it may be the case that for some effort levels $e > 1$ the IR constraint is binding instead of the LL constraint, and the principal need not give up any limited-liability rent. Fourth, we extend the proof of Theorem 3.1 to account for our new characterization of the minimum expected payments.

**Step 1.** Fix an effort level $1 < e < E$. We need to show that $m_e \leq m_{e+1}$. Because the effort level $e$ is implementable, there exists a wage schedule $(w^1, w^2)$ satisfying the IC constraints, which in the case of two outcomes reduce to $(\pi^2_e - \pi^2_f)(w^2 - w^1) \geq c_e - c_f$ for $f \neq e$. The IC constraint for effort level $e-1$ together with FOSD implies that $w^2 - w^1 \geq m_e$. Similarly
the IC constraint for effort level $e + 1$ gives that $w^2 - w^1 \leq m_{e+1}$. The result follows from combining the last two equations.

**Step 2.** Take any two consecutive implementable effort levels $e$ and $e + 1$, such that $u_{e}^{SW} \leq u_{e+1}^{SW}$ and $1 < e < E$. To conclude that the sequence of social-welfare values is quasi-concave and satisfies QCSW, we will show that effort level $e - 1$ verifies $u_{e-1}^{SW} \leq u_{e}^{SW}$. Using again FOSD and that there are two outcomes, we can rewrite the condition $u_{e}^{SW} \leq u_{e+1}^{SW}$ as $m_{e} \leq y^{2} - y^{1}$. Now R-IMCP implies that $m_{e} \leq m_{e+1} \leq y^{2} - y^{1}$, and we obtain that $u_{e}^{SW} \leq u_{e}^{SW}$ as required.

**Step 3.** We will show that minimum expected payments that the principal makes to the agent are

$$z_1 = c_1, \quad \text{and} \quad z_e = \max\left(c_e, \pi_e^2 m_e\right) \quad \text{for } 1 < e \leq E. \quad (9)$$

The result follows trivially for effort level $e = 1$ by considering the wage schedule at the intersection of the IR constraint and $w^2 = 0$. Fix an effort level $1 < e \leq E$. Because R-IMCP imposes no condition for the lowest effort level we need to consider two cases when characterizing the optimal solution of MPLP$_e$: whether $c_e \leq \pi_e^2 m_e$ or $c_e > \pi_e^2 m_e$. The former case guarantees that the LL constraint is binding, the analysis of Proposition 2.1 applies and $z_e = \pi_e^S m_e$. In the latter case we need to show that the IR constraint is binding instead, the principal gives up no limited-liability rent, and $z_e = c_e$.

Assume that $c_e > \pi_e^2 m_e$. We shall show that the wage schedule lying at the intersection of the IR constraint and the IC constraint of effort level $e - 1$ is optimal. Some algebra shows that this wage schedule is given by $w^1 = (\pi_{e-1}^2 c_{e-1} - \pi_{e-1}^2 c_e) / (\pi_e^2 - \pi_{e-1}^2)$ and $w^2 = (\pi_{e-1}^1 c_e - \pi_{e-1}^1 c_{e-1}) / (\pi_e^2 - \pi_{e-1}^2)$. Because this wage schedule leaves no utility for the agent, feasibility of the wage schedule suffices to prove its optimality. For the LL constraint, note that $w^2 \geq 0$ from FOSD and the fact that costs are non-decreasing. For $w^1 \geq 0$ we need that $\pi_e^2 c_{e-1} \geq \pi_{e-1}^2 c_e$, which follows after some algebra from our condition that $c_e > \pi_e^2 m_e$. For the IC constraints note that $w^2 - w^1 = m_e$, and thus these reduce to showing that $m_e \geq m_{e,f}$ for $1 \leq f < e$ and $m_e \leq m_{e,f}$ for $e < f \leq E$. The latter holds by R-IMCP and the steps in the proof of Lemma A.1 restricted to effort levels $1 < e \leq E$.

**Step 4.** We conclude by extending the proof of Theorem 3.1 to account for our new characterization of the minimum expected payments in (9). As before applying Corollary 2.2 (which can be easily shown to hold under (9)) the principal has no incentive to implement an effort level higher than the largest socially-optimal effort level $e^{SO}$.
Let $\tilde{e} \triangleq \max\{e \in E : z_e = c_e\}$ be the greatest effort level that leaves no rent to the agent. Clearly, such an effort level exists since the lowest effort level $e = 1$ satisfies that property. When $\tilde{e} > e^{SO}$ Corollary 2.2 implies that the optimal effort level for the principal is the socially-optimal one and the welfare loss is one. In the remainder of the proof we consider the case when $\tilde{e} \leq e^{SO}$. Notice further that the second-best effort level satisfies $e^* \geq \tilde{e}$ since effort levels lower that $\tilde{e}$ provide zero utility for the agent and QCSW implies that the principal’s utility for lower effort levels is dominated by that of $\tilde{e}$. The proof follows from repeating the steps of the proof of Theorem 3.1 for effort levels $e$ such that $e^* < e \leq e^{SO}$, for which the optimal payment is $z_e = \pi^S_em_e$ as before.

\section*{A.6 Proof of Theorem 4.1}

\textbf{Proof.} We prove the result in four steps. First, we take a principal-agent problem $I = (\pi, y, c) \in \mathbb{I}^{E,S}$ restricted to linear contracts with $E$ effort levels and $S$ outcomes, and construct an unrestricted reduced problem $\tilde{I} = (\tilde{\pi}, \tilde{y}, c) \in \mathbb{I}^{E,2}$ with $E$ effort levels and 2 outcomes. Second, we show that for each effort level $e \in E$ the linear program MPLP$_e$ for instance $I$ restricted to linear contracts is equivalent to the linear program MPLP$_e$ for the reduced instance $\tilde{I}$. Third, we show that the reduced instance $\tilde{I}$ satisfies assumptions FOSD and IMP, and $\tilde{y}$ is non-negative. Finally, we conclude by showing that the bound holds.

\textbf{Step 1.} In showing the reduction, we write the contracts in terms of wage at the lowest outcome $d = a + by^1$, and consider the wages $w^s = d + b(y^s - y^1)$, which is valid because the parameters are not restricted. Then the reduced instance is given by

\begin{align*}
\tilde{\pi}_e &\triangleq (1, \pi_e y - y^1), \quad \forall e \in E, \\
\tilde{y} &\triangleq (y^1, 1),
\end{align*}

and additionally we refer to the reduced wage schedule as $\tilde{w} \triangleq (d, b)$. The first entry in the reduced wage vector $\tilde{w}$ represents the wage at the lowest outcome while the second represents the slope of the linear contract. The expected payment under effort level $e \in E$ is $\pi_e w = d + b(\pi_e y - y^1) = \tilde{\pi}_e \tilde{w}$, and the expected output satisfies $\pi_e y = \tilde{\pi}_e \tilde{y}$, as expected.

Notice that the reduced vectors $\tilde{\pi}_e$ for $e \in E$ no longer sum up to one, and thus are not probability distributions. Surprisingly, this will not be important for our results. Notice that the reduced output may not be increasing, but again our results will hold in this case too. By interpreting the entries of the reduced vectors above as two outcomes, we show that the original problem and the reduced problem are equivalent.
**Step 2.** Fix an effort level $e \in E$. For the second point, under linear contracts, the original minimum payment linear program corresponding to effort level $e$, $\text{MPLP}^L_e$ can be written as

$$z_e = \min_{a,b} a + b\pi_e y$$

s.t. $a + b\pi_e y - c_e \geq 0$ \hspace{1cm} (10a)

$b\pi_e y - c_e \geq b\pi_f y - c_f \hspace{1cm} \forall f \in E \setminus e$ \hspace{1cm} (10b)

$a + by^s \geq 0 \hspace{1cm} \forall s \in S$. \hspace{1cm} (10c)

We first show that given a feasible solution $(a, b)$ of $\text{MPLP}^L_e$, then the reduced solution $\tilde{w} = (d, b)$ with $d = a + by^1$ is feasible for $\text{MPLP}_e$ of the reduced instance $\tilde{I}$ and achieves the same objective. We have seen in step 1 that the objective is preserved by the reduction. The original individual rationality constraint (10a) can now be written as $\tilde{\pi}_e \tilde{w} \geq c_e$, while the original incentive compatibility constraints (10b) can be written as $\tilde{\pi}_e \tilde{w} - c_e \geq \tilde{\pi}_f \tilde{w} - c_f \forall f \neq e$. For outcome $s = 1$ the original limited liability constraint (10c) implies that $d \geq 0$. For an effort level $e > 1$ we have by the incentive compatibility constraint (10b) with effort level $f < e$ that $b \geq (c_e - c_f)/(\pi_e y - \pi_f y) \geq 0$, where the last inequality follows from FOSD. For the lowest effort level $e = 1$, one always has that $d = c_1$ and $b = 0$ is an optimal solution. Hence, the constraint $b \geq 0$ does not eliminate any optimal solution. This completes this direction of the reduction.

For the opposite direction, we need to show that given a feasible solution $\tilde{w} = (d, b)$ for $\text{MPLP}_e$ of the reduced instance $\tilde{I}$, the solution $(a, b)$ with $a = d - by^1$ is feasible for $\text{MPLP}^L_e$ and achieves the same objective. Notice that (10a) and (10b) follow directly from the reduced problem’s IR and IC constraints respectively. From the reduced problem’s LL we know that $b \geq 0$ and $d \geq 0$. Hence for all $s \in S$ we have that $a + by^s \geq a + by^1 = d \geq 0$, because the outputs are non-decreasing. Thus, (10c) holds and the reduction is complete.

**Step 3.** First, we prove that $\tilde{\pi}$ satisfies FOSD. We require that $\tilde{\pi}_e^2 \leq \tilde{\pi}_f^2$ for $e \leq f$. The previous is equivalent to $\pi_e y \leq \pi_f y$ that holds because the original distribution $\pi_e$ satisfies FOSD and the output is non-decreasing. A similar argument shows that $\tilde{\pi}_e \geq 0$. Finally, the equivalence given in step 2 shows that an effort level $e \in E$ is implementable in $\text{MPLP}^L_e$ for instance $I$ if and only if it is implementable in $\text{MPLP}_e$ for the reduced instance $\tilde{I}$. This shows that the reduced instance $\tilde{I}$ satisfies IMP.

**Step 4.** Finally, notice that in the proof of Theorem 3.2 we do not use the fact $\tilde{y}$ is increasing, and the result still holds if $\tilde{\pi}_e$ do not sum up to one. Thus all our results apply
to the reduced problem. Putting everything together we obtain that

$$
\rho^L(I) = \rho(\tilde{I}) \leq 1 + \sum_{e=2}^E \frac{\pi_e Y - \pi_{e-1} Y}{\pi_e Y - y^1},
$$

where the first equality follows from the reduction in step 2, and the inequality from Theorem 3.2.

A.7 Proof of Lemma 4.2

Proof. Let \( \{X_e\}_{e \in E} \) be a family of random variables, such that \( X_e \) is the random number of successes given that the agent works hard in \( e \) tasks. Then, \( X_e \) is the sum of \( e \) independent Bernoulli random variables with success probability \( p_h \), and \( N - e \) independent Bernoulli random variables with success probability \( p_l \). Denote by \( Y(p) \) a Bernoulli random variable with success probability \( p \); i.e., \( \mathbb{P}(Y(p) = 1) = p = 1 - \mathbb{P}(Y(p) = 0) \). Hence, we may write \( X_e \) as

$$
X_e = \sum_{f=1}^e Y_f(p_h) + \sum_{f=e+1}^N Y_f(p_l) = \sum_{f=1}^N Y_f(p_f(e)),
$$

where the functions \( \{p_f(e)\}_{f \in E} \) equal \( p_l \) if \( e < f \), and \( p_h \) otherwise. Notice that for all \( f \in E \) the functions \( p_f(e) \) are non-decreasing in \( e \). Ghurye and Wallace (1959) or more recently Huynh (1994) show that given any number of independent Bernoulli random variables \( Y_f \) with success probability \( p_f(e) \) strictly between 0 and 1, and non-decreasing in \( e \), then the sum \( \sum Y_f \) has monotone likelihood-ratio with respect to \( e \).

A.8 Proof of Theorem 4.3

Proof. First, Laux (2001) shows that in the case of multiple tasks the principal either implements the highest or the lowest aggregate effort levels. The minimum payments are given by \( z_0 = N c_l \) for the lowest effort level, and

$$
z_N = \max \left( N c_h, N p_h^N \frac{c_h - c_l}{p_h^N - p_l^N} \right),
$$

for the highest effort level. Note that Corollary 2.2 applies and the agent’s rent is non-decreasing with the effort level. Indeed, using that \( c_N = N c_h \) and that \( c_0 = N c_l \) we obtain that \( z_N - c_N \geq 0 = z_0 - c_0 \) as expected.

Second, we show that the social welfare is linear in the number of aggregated effort levels and thus satisfies the QCSW assumption. Referring to the number of successes given
that the agent works hard in $e$ tasks by the random variable $X_e$, its expected number is $E[X_e] = ep_h + (N - e)p_l$. One may write the social welfare as

$$w^{sw}_e = \pi e y - c_e = E[X_e \bar{y} + (N - X_e)\bar{y}] - c_h e - c_l(N - e)$$

$$= N \left[p_l \bar{y} + (1 - p_l)\bar{y} - c_l \right] + e \left[(p_h - p_l)(\bar{y} - y) - (c_h - c_l) \right],$$

which is linear in effort level $e$. The first term can be interpreted as the baseline social welfare when the agent exerts a low effort level in all tasks, while the second term can be interpreted as the marginal contribution to social welfare of each additional task for which the agent exerts high effort level.

Third, because social welfare is linear, the largest socially-optimal effort level is either $e^{SO} = 0$ or $e^{SO} = N$. The principal has no incentive to implement any effort level higher than $e^{SO}$, because the utility of the agent is non-decreasing with his effort level. In the case when $e^{SO} = 0$ we conclude that the optimal effort level for the principal is $e^* = 0$, the welfare loss is one, and the bound trivially holds. In the remainder of the proof we shall consider the case when $e^{SO} = N$. If aggregate effort level $N$ leaves no rent to the agent, or equivalently $z_N = c_N$, then it is optimal for the principal to induce the high effort level $e^* = N$. Inefficiency is thus introduced when the principal induces the low effort level $e^* = 0$, and $z_N = Np^N_h(c_h - c_l)/(p^N_h - p^N_l)$. In this case we apply the last steps of the proof of Theorem 3.1 to obtain inequalities (7) and (8), and conclude that the welfare loss is upper bounded by

$$1 + \frac{p^N_l - p^N_h}{p^N_h} = 2 - \left(\frac{p_l}{p_h}\right)^N < 2.$$

This bound is tight.