

OPPORTUNISM IN PRINCIPAL-AGENT RELATIONSHIPS WITH SUBJECTIVE EVALUATION

W. BENTLEY MACLEOD AND TECK YONG TAN

ABSTRACT. We consider a general principal-agent model with subjective evaluation and show that order of information revelation in such contracts has sharp predictions regarding both contract form and its efficiency. This perspective allows one to use agency theory to define and explore two large classes of contracts – authority contracts where compensation is set by the Principal, and sales contracts where payment is set by the Agent. We show that sales contracts provide a way to model Williamson’s (1975) notion of *opportunism* (self-interest seeking with guile), where guile is interpreted as the *combined* risk of the agent shirking and lying. The approach provides a way to precisely model “good faith” in contract.

1. INTRODUCTION

An open challenge for contract theory is to explain the wide variety of observed contracting relationships. In a classic paper, [Kerr \(1975\)](#) provides a number of examples of organizations that introduced dysfunctional performance pay systems to illustrate that even successful firms are often challenged by the problem of incentive pay design. Moreover, even though legions of business school students are taught this paper and despite the input of many highly trained individuals, we still observe today many examples of poorly designed reward systems, particularly in the areas of health care ([Frank and McGuire, 2000](#); [Skinner, 2012](#)) and education ([Goodman and Turner, 2010](#)).

Date: February 25, 2017.

Key words and phrases. Subjective Evaluation, Guile, Opportunism, Sales Contracts, Authority Contracts, Malfeasance.

We thank Yeon-Koo Che, Bob Gibbons, Bengt Holmström, Navin Kartik, Lewis Kornhauser, Tracy Lewis, Jim Malcomson, Stephen Morris, Joel Sobel and seminar participants at MIT, the Conference of Society for Institutional & Organizational Economics 2016, the Econometric Society meetings 2016 at San Francisco, U-Penn, Annual Conference on Relational Contracts, Madrid 2016, and Queen’s University for helpful comments. We also thank Elliot Ash for research assistance.

One common difficulty faced in designing good compensation systems is the lack of quality information regarding performance.¹ A solution to this problem is to rely upon subjective measures of performance instead. In this paper, we show that the form of the optimal contract is very sensitive to how subjective evaluations by the Principal (he) and Agent (she) are incorporated. Specifically, holding the production environment fixed, we show that the order of information revelation has sharp and distinct predictions regarding the contract form. Moreover, there can be incentive for a Principal in some cases to offer contracts that may lead to the opportunistic behaviors observed by [Kerr \(1975\)](#).

Canonical contracting models suppose that parties have a good understanding of the environment and would offer contracts that are optimal conditional upon the characteristics of the environment. In the classic principal-agent model, the Principal offers the risk-averse Agent a compensation that is a function of the commonly observed performance. Given that performance measures are necessarily imperfect, the optimal contract trades off the insurance benefit from more rigid wages against the benefit of increasing effort incentives by increasing the sensitivity of pay to performance. This leads to the informativeness principle which predicts that pay should vary with any available information that concerns the Agent's performance, and hence compensation should in general be sensitive to any piece of relevant information regarding performance.

Although agency theory has been successful in explaining the structure of observed compensation contracts as well as features of agricultural share-cropping contracts, the price terms in observed contracts are often not as sensitive to information held by contracting parties as the theory predicts. A goal of this paper is to show that the contract form is not only sensitive to the quality of information, but also to the timing of information revelation. This allows us to explain some features of observed contracts that cannot be explained with the canonical principal-agent model ([MacLeod and Parent 1999](#); [Prendergast 1999](#); [Gibbons 2005](#)).

The importance of the timing of information release is first studied in the classic paper by [Holmström and Milgrom \(1987\)](#). They show that when actions are taken over time and rewards are based upon the entire path of performances, it is optimal to use a linear reward system that ensures a constant marginal return to effort at each point in time. [Baker \(1992\)](#) introduces a model in which effort is chosen after the Agent receives further information about how her effort would affect the measured performance. He shows that this can lead to the firm reducing the sensitivity of the Agent's pay to performance in order to decrease the

¹As [MacLeod and Parent \(1999\)](#) document, workers are more likely to be given performance pay when quality measures of performance exist, as they do in sales and machine tool production (see Tables 3a and 3b in the paper).

incentive for the Agent to “game” the compensation system. [Prendergast \(2002\)](#) extends this work to show that the informative principle is not general, but also depends upon the timing of information release. In particular, the Agent may face greater sensitivity of pay to performance in riskier environments if the Agent has better information that can be used to mitigate risk. Conversely, compensation places more weight upon input measures when there is more idiosyncratic risk that cannot be controlled by the Agent.

In this paper, we explore the effects of timing of information release upon the contract form when performance evaluations are subjective and verifiable measures of performance are not available. The early literature in relational contract, beginning with [Telser \(1980\)](#), [Bull \(1987\)](#), and [MacLeod and Malcomson \(1989\)](#), supposes that the Principal and the Agent have common knowledge regarding performance but cannot enforce contracts based on it. In this case, contract enforcement relies upon a combination of gains from future trade and coordinated punishments for bad behaviors by either party. [Baker et al. \(1994\)](#) and [Schmidt and Schnitzer \(1995\)](#) observe that the combination of subjective and verifiable signals can in some cases render relational contracts unenforceable.² This can lead to the result that more information undermines the operation of markets, a point that has been made in the context of market trade by [Kranton \(1996\)](#) and [Sobel \(2006\)](#).

[Levin \(2003\)](#) introduces a model of relational contract with subjective evaluation in which the Principal privately observes the Agent’s performance, and he shows that one can implement efficient effort with a one-step contract that punishes the Agent when the Principal’s subjective evaluation falls below a threshold. The contract is incentive compatible because the Principal is indifferent being punishing and not punishing the Agent. Hence it is technically identical to the efficiency wage contract of [Shapiro and Stiglitz \(1984\)](#), where the firm in [Shapiro and Stiglitz \(1984\)](#) is also indifferent between keeping a worker or replacing her with a new worker; hence it is optimal for the firm to fire a worker if and only if the firm perceives that the worker has shirked, which in turn incentivizes the worker to work because job loss is costly to her.

[MacLeod \(2003\)](#) extends this analysis to a more general asymmetric information structure. Modeling relational contracts in a repeated game framework makes the analysis of equilibria under more complex information structures extremely challenging – see [Kandori \(2002\)](#) for a discussion and [Kandori \(2011\)](#) for some recent results. [Levin \(2003\)](#) circumvents this problem by restricting attention to “full-review” contracts which ensure that parties play with common knowledge regarding future payoffs at the beginning of each

²Relatedly, [Zabojnik \(2014\)](#) illustrates how objective measures of performance helps to sustain relational contracts on subjective evaluations when the subjective evaluations also have to play a feedback role to the Agent.

period, a case that has been relaxed by [Fuchs \(2007\)](#) and [Malcomson \(2016\)](#). [MacLeod \(2003\)](#) addresses this problem by relying upon a reduced form model of relational contract where the Agent has the capability to impose conflict costs upon the Principal, and the optimal contract modulates this conflict as a function of the information that the players have. As an empirical matter, one only needs to look at the daily papers to see the many examples of individuals imposing conflict costs upon others. The notion that these costs can be strategic is due to [Frank \(1987, 1988\)](#). It is also well-known in experimental economics that individuals may inflict costly punishments on other parties when they perceive themselves to have been treated unfairly (e.g. [Guth et al., 1982](#) and [Levine, 1998](#)), and [Mas \(2006, 2008\)](#) provide direct evidence in an employment context of employees imposing costs upon employers through private actions.

[MacLeod's \(2003\)](#) main contribution is to show that the Agent's ability to harm the Principal can be an essential input into an optimal contract with subjective evaluation. In contrast to the model of [Levin \(2003\)](#) where only the Principal is informed, [MacLeod \(2003\)](#) allows for both parties to have information and shows that higher correlation in the parties' information reduces the expected level of conflicts in an optimal contract.³ This result is consistent with the message from compensation experts who emphasize the importance of transparency and ensuring employees understand and accept the compensation package offered by the firm ([Milkovich et al., 2011](#)).

A key ingredient in [MacLeod \(2003\)](#) is the assumption that the Principal reveals his information before the Agent does.⁴ As a result, the Agent's information cannot be used to set compensation even if it is useful. This is in line with a common assumption in agency theory that the Principal is the most informed party regarding performance, so it is natural to suppose that the Principal uses his information to set compensation. However, there are many situations where the Agent is likely to be more informed. For example, the Agent might be a physician, stock broker or car mechanic whom the Principal seeks advice from. Advice is often a credence good – a good or service whose quality is not known until after the purchase. In such situations, the Agent is often the one who determines prices as a function of her information, and the market for such goods has typically been studied as a problem that is distinct from the principal-agent problem ([Darby and Karni, 1973](#); [Emons, 1997](#); [Dulleck and Kerschbamer, 2006](#)).

³[Chan and Zheng \(2011\)](#) and [Maestri \(2012\)](#) also allow the Agent to have information about performance but they consider a more specific information structure.

⁴Proposition 3 in the published version of [MacLeod \(2003\)](#) is wrong, as first pointed out by Alexander Nekrasov. All the results in [MacLeod \(2003\)](#) hold for what is called an “authority contract” in this paper. Proposition 3 in [MacLeod \(2003\)](#) should be replaced by Lemma 3 in this paper.

In this paper, we focus upon the case in which both the Principal and Agent are risk-neutral and have private information about the performance. The goal of the contract is to design rewards based on both parties' information that minimizes expected conflict costs while inducing effort from the Agent. Holding the production environment fixed, we consider two distinct contractual arrangements that differ in their order of information revelation:

- (1) Under an *authority contract*, the Principal reveals his subjective evaluation regarding the Agent's performance first, and then the Agent reveals hers.
- (2) Under a *sales contract*, the Agent reveals her subjective evaluation regarding her performance first, and then the Principal reveals his.

We find that the order of information revelation has striking implications for the contract form. Specifically, the party who moves last has his/her rewards being insensitive to his/her information. Hence the resulting contract form of an authority contract resembles an employment contract where the Agent's compensation is determined by the Principal's information, while a sales contract resembles a contract for credence goods where the Principal's payment is determined by the Agent's information.

Our model also provides a way to formally integrate Oliver Williamson's notion of *opportunism* into contract theory:

“Opportunism extends the conventional assumption of self-interest seeking with guile and has profound implications for choosing between alternative contractual relationships.” - [Williamson \(1975\)](#), page 26.

We view Williamson's notion of “self-interest” as the need for a contract to ensure that there is no gain from one-shot deviation by either party. More precisely, the contract must ensure that it is in the Agent's best interest to exert the effort obligation when she believes that both parties will be truthful in revealing their private information later. Similarly, under the hypothesis that the Agent has exerted the effort obligation, the contract gives incentives for both parties to truthfully reveal their private information.

This view of self-interest is in line with [Levin \(2003\)](#) who uses the term “good faith” to describe incentive compatible relational contracts. Building upon the insights of [Abreu \(1988\)](#), [Levin \(2003\)](#) shows that the strategies constitute perfect Bayesian equilibria when neither party can gain from a one-shot deviation. In our model, good-faith contracts (i.e. contracts that players cannot gain from a one-shot deviation) are perfect Bayesian equilibria for authority contracts, but not necessarily so for sales contracts. In particular, the Agent can possibly game a good-faith contract via a “double deviation” in which the Agent consciously deviates from the effort obligation and follows it up with a misrepresentation

of her private information to cover up her tracks.⁵ Such forms of strategic manipulation corresponds nicely to Williamson’s requirement that opportunism entails “guile” – the Agent strategically modifies effort and then lies about it.

We say that a contract is “guile-free” if there is no incentive to engage in opportunistic behavior – a strategic manipulation. Guile-free contracts by construction implement actions that form perfect Bayesian equilibria. We show that ensuring good-faith contracts to be also guile-free requires adding a number of new incentive constraints that increases exponentially with the signal space. This has a number of interesting implications, or (in the language of Williamson) profound implications for choosing between alternative contractual relationships.

First, the Principal’s expected costs under a good-faith contract is lower than that of a guile-free contract. Given that opportunistic behavior requires complex strategic manipulation, it can be optimal for a Principal to then offer a good-faith sales contract should he believe that the Agent will not be opportunistic. Moreover, this continues to be an optimal choice if opportunism by Agents is not so common. At the same time, the existence of some opportunistic Agents implies that we may observe contracts being gamed in equilibrium, a result that is not predicted by standard agency models, but as we have mentioned, is a feature of some observed contractual relationships.

Second, if opportunism is a binding constraint that significantly reduces the effectiveness of a sales contract, then it may be more efficient for parties to use an authority contract. Such a contract can naturally be viewed as representing an employment relationship, and hence the potential for opportunism may explain why some relationships are brought inside the firm. This result applies to contracts where parties must rely upon subjective evaluations for rewards. Conversely, if performance becomes more easily verifiable, then the reverse can occur. For example, easier monitoring of drivers makes it possible for more arm’s length relationships with limousine drivers, as we have seen with the rise in popularity of Uber, a firm that has developed a number of observable performance metrics for the both drivers and their passengers (Cramer and Krueger, 2016).

The agenda of the paper is as follows. Section 2 introduces the model together with the signal-generating process that forms the basis of the parties’ subjective performance evaluations. Section 3 considers the sales contracts and introduces the notion of guile. We introduce an example illustrating a case in which there is no authority contract implementing positive effort, even though a feasible sales contract does exist. Section 4 discusses when opportunism is a binding constraint and provides a number of illustrative examples.

⁵See Deb et al. (2016) for a discussion of double deviations in a model with peer effects.

Section 4.3 considers an “expert-agent” environment to illustrate the importance of the underlying information structure on the presence of opportunism. Section 5 introduces the authority contract. We conclude with a discussion in Section 6. Omitted proofs are found in Appendix A.

2. MODEL

2.1. The Environment. Consider a static Principal-Agent model with risk-neutral parties.⁶ The Agent has unlimited liability and an outside option of U^0 , and his effort is privately chosen from $\lambda \in [0, 1]$. Performance, or synonymously, outcome is binary: $o \in \{L, H\}$. Under λ , the probability of getting H is λ and the probability of getting L is $1 - \lambda$. H represents a high performance that generates revenue $B > 0$ to the Principal. On the other hand, L is a low performance that generates no revenue. The Agent’s effort cost function is $V(\cdot)$ which satisfies:

Assumption 1. $V : [0, 1) \rightarrow \mathfrak{R}_+$ is any twice differentiable, strictly increasing and strictly convex function such that $V(0) = 0$, $V'(0) > B$, with $V(1) = \lim_{\lambda \rightarrow 1} V(\lambda)$ when the limit exists.

Let λ^{FB} be the efficient effort level. If $B \geq V'(1)$, then $\lambda^{FB} = 1$; otherwise it is uniquely defined by $V'(\lambda^{FB}) = B$. Hence the first best surplus is strictly positive:

$$(2.1) \quad \text{Surplus}^{FB} = \lambda^{FB}B - V(\lambda^{FB}) > 0.$$

The outcome H or L is not directly observable.⁷ Instead, the Principal and the Agent each receives a *private* signal about the outcome,⁸ denoted respectively by $t, s \in S = \{0, 1, \dots, n - 1\}$, where S is a finite set with $n \geq 2$ elements. These signals have the same “meaning” to the two parties. For example, if $s = 10$ means that the performance is excellent according to the Agent, the Principal will have the same interpretation for $t = 10$. However, he may perceive the quality differently when he receives a signal of, say, $t = 8$ which means that the performance is good but not excellent. The two parties can then reasonably disagree on the performance level. This provides a precise notion of subjective evaluation where the extent to which the parties’ signals are correlated provides a measure of the subjectivity level (or the lack of it); when signals are perfectly correlated, we are in

⁶We use the male pronoun for the Principal and the female pronoun for the Agent.

⁷For concreteness, one can think of B as some future payoff in which the value is not immediately realized. For example, the product might be a creative output such as painting or screenplay, the quality of which increases with effort, but its value is only determined after several years.

⁸This asymmetry in information on the outcome prohibits the Principal from efficiently “selling the firm” to the Agent.

the case of common knowledge of performance as in the literature on relational contracts (Telser, 1980; Bull, 1987; MacLeod and Malcomson, 1989).

The signal-generating process begins with a set of states given by $ts \in S^2$. We refer to t observed by the Principal or s observed by the Agent as a *signal*, and the joint-realization ts as a *state*. Let the probability of getting state ts under outcome $o \in \{L, H\}$ be given by

$$\Pr[ts|o] = \Gamma_{ts}^o.$$

If the Principal correctly anticipates the effort λ exerted by the Agent, both parties will have the same *ex-ante* unconditional probability of state ts :

$$\begin{aligned} \Gamma_{ts}(\lambda) &= \lambda \Gamma_{ts}^H + (1 - \lambda) \Gamma_{ts}^L \\ &= \Gamma_{ts}^L + \lambda \hat{\Gamma}_{ts}, \end{aligned}$$

where $\hat{\Gamma}_{ts} = \Gamma_{ts}^H - \Gamma_{ts}^L$ is the marginal effect of effort λ on the probability of state ts .

It will be useful to exploit the linear structure of the problem by representing the probability distribution as a $1 \times n^2$ row vector. We take the convention of lexicographically ordering states ts by the Agent's signal s first and then the Principal's signal t . Hence $\vec{\Gamma}^o$ is the probability vector for outcome o with:⁹

$$\vec{\Gamma}^o = \left[\underbrace{\Gamma_{00}^o, \Gamma_{10}^o, \dots, \Gamma_{(n-1)0}^o}_{s=0}, \underbrace{\Gamma_{01}^o, \Gamma_{11}^o, \dots, \Gamma_{(n-1)1}^o}_{s=1}, \dots, \underbrace{\Gamma_{0(n-1)}^o, \Gamma_{1(n-1)}^o, \dots, \Gamma_{(n-1)(n-1)}^o}_{s=n-1} \right].$$

We represent vector $\vec{\Gamma}$ analogously for the $1 \times n^2$ vector of $\hat{\Gamma}_{ts}$, and the unconditional probability vector of state can be written as:

$$\vec{\Gamma}(\lambda) = \vec{\Gamma}^L + \lambda \vec{\Gamma} \in \mathfrak{R}^{n^2}.$$

2.2. Contracts. By the revelation principle, it is sufficient to consider only contracts with the feature that parties are truthful in equilibrium. For example, Levin (2003) formally allows the Principal to send a message, which in turn is a function of the Principal's private information. As is common in the literature, we subsume the message game in the relationship between information and the contract. Hence a contract in this model specifies the costs for the Principal and wages for the Agent under each reported state.

In addition, we allow both the Principal and the Agent to costlessly impose a deadweight loss upon each other. This captures the notion of conflict in a relationship. Myerson and

⁹Formally, we index each state by an index function $\mathcal{I} : S^2 \rightarrow \{0, 1, \dots, n^2 - 1\}$ where $\mathcal{I}(ts) = (n \times s) + t$. Thus for vector $\vec{x} \in \mathfrak{R}^{n^2}$, x_{ts} refers to the $\mathcal{I}(ts)$ -th entry of vector \vec{x} .

Satterthwaite (1983) and Crampton (1985) observe that trade under asymmetric information necessarily entails social loss to ensure parties reveal their private information. The extent to which such conflicts are possible and required in equilibrium depends upon the context. For example, if the Principal feels over-charged by the Agent, he might retaliate by harming the Agent’s reputation on social media.¹⁰ Alternatively, if the Agent feels under-compensated by the Principal, she might retaliate with lower-quality services in the future or likewise try to harm the reputation of the Principal. In this sense, such conflicts can be viewed as “behavioral” responses that are implemented when the parties disagree on their evaluations of the performance level.¹¹

Accordingly, let c_{ts} be the Principal’s cost and w_{ts} be the Agent’s wage when the reported state is ts . These costs and wages are net of the conflict imposed by the other party and satisfy the *relaxed budget constraint (RBC)*:

$$(2.2) \quad c_{ts} \geq w_{ts}, \forall ts \in S^2.$$

The social deadweight loss due to the conflict at state ts is:

$$(2.3) \quad \delta_{ts} = c_{ts} - w_{ts} \geq 0.$$

This social loss is the sum of conflicts imposed by the Principal upon the Agent, and by the Agent upon the Principal. What is crucial is that these conflicts are pure losses; they are not transfers.¹² Given that parties care only about their wage w_{ts} or cost c_{ts} , it follows that *who* imposes the conflict is indeterminate and thus we can choose who to impose the conflict in a way that is more convenient for the analysis. This follows from:

Lemma 1. *In the absence of constraints on the size of the conflict δ , it is without loss of generality to consider contracts where only one party (the Principal or the Agent) is inflicting the conflict.*

Analogous to the vector representation for the probability distribution of states, we denote \vec{c} , \vec{w} and $\vec{\delta}$ as the vector representation of the cost, wage and conflicts terms, with the same ordering as previously (i.e. s first and then t lexicographically). We take the

¹⁰Banerjee and Duflo (2000) and Macchiavello and Morjaria (2015) provide evidence on the value of reputation in relationships with repeated interactions.

¹¹One could also view these conflicts as the level of aggrievement in the sense of Hart and Moore (2007).

¹²In some cases, employment can be organized so that these losses are implemented as a transfer to other workers (see Malcomson (1984) and Carmichael (1983)). In this case, these losses are then no longer social losses. However, doing so introduces the problem of collusion among Agents and contracts will then have to be designed to be collusion-proof, which adds further complexity to the contract design.

convention that these are $n^2 \times 1$ *column* vectors so that for $x \in \{c, w, \delta\}$, we can write the expectation of x as the inner product: $\vec{\Gamma}(\lambda) \vec{x} = \sum_{ts \in S^2} \Gamma_{ts}(\lambda) x_{ts}$.¹³

It is shown later that the problem is convex and hence there is no gain from randomization. Hence a contract implementing effort $\lambda \in [0, 1]$ is defined by:

$$(2.4) \quad \psi = \{\lambda, \vec{c}, \vec{w}\} \in \Psi \equiv [0, 1] \times \mathbb{R}^{n^2} \times \mathbb{R}^{n^2},$$

which specifies the Agent's effort obligation λ , together with the Principal's cost and the Agent's wage at each reported state under the restriction of RBC (2.2). The conflict terms in the contract are then determined by $\vec{\delta} = \vec{c} - \vec{w}$. Effort λ is not directly contractible since it is privately exerted by the Agent, so the wage terms \vec{w} have to provide incentives for the Agent to adhere to the effort obligation.¹⁴

With regards to information revelation, we consider first the *sales relationship* in which the Agent makes a report on her private signal s first and then the Principal, upon observing the Agent's report, makes a report on his signal t . As we discussed in the introduction, the timing is intended for situations where the Agent is likely to have superior information to the Principal, such as in the provision of consulting services. The heuristic sequence of moves for the contracting game is as follows:

¹³All multiplications between vectors in this paper refer to the inner product.

¹⁴In defining the contract this way, we are implicitly assuming that the players are contracting upon a joint payment of $\delta_{ts} = c_{ts} - w_{ts}$ to a third party at each reported state ts . This interpretation is convenient from a modeling perspective, but it might not be obvious to readers that this also fits into our interpretation of δ as a reduced-form punishment in the relationship, especially when it might be difficult to imagine that the level of conflict can be contracted upon. To reconcile this, one can take the interpretation that the only contractible term in the contract is a price term p_{ts} that the Principal must pay to the Agent at each reported state ts . On top of that, the contract also specifies a set of punishment *recommendations* (d_{ts}^A, d_{ts}^P) for each reported state ts . d_{ts}^A is the recommended conflict cost that the Agent should be imposing upon the Principal in reported state ts , and likewise d_{ts}^P is the recommended conflict cost that the Principal should be imposing upon the Agent. Like effort λ , the punishment recommendations have to be incentive compatible in the sense that it is the best response of the players to obey these punishment recommendations. But since it is costless to the inflicting party to inflict conflict costs on the other party, obeying the punishment recommendations will always be a best response. The Principal's cost term in the contract in (2.4) is then given by $c_{ts} = p_{ts} + d_{ts}^A$ and the Agent's wage term is $w_{ts} = p_{ts} - d_{ts}^P$.

TABLE 2.1. Action Timing

Step	Actor	Actions
1.	Principal	Offers a contract $\psi \in \Psi$ to the Agent.
2.	Agent	“Accept” or “Reject”.
		If “Reject”, relationship ends; if “Accept”, go to step 3.
3.	Agent	Privately select effort $\lambda \in [0, 1]$.
4.	Nature	The outcome $o \in \{L, H\}$ is realized, which in turn generates private signals t and s for the Principal and the Agent respectively.
5.	Agent	Reports \hat{s} regarding her signal s .
6.	Principal	Observes \hat{s} and reports \hat{t} regarding his signal t .
7.	Court	Principal pays $c_{\hat{t}\hat{s}}$ and Agent receives $w_{\hat{t}\hat{s}}$.

When the parties abide by the conditions of the contract and report their signals truthfully, the expected payoffs of the Principal and the Agent are respectively:

$$(2.5) \quad U^P(\psi) = \lambda B - \vec{\Gamma}(\lambda) \vec{c},$$

$$(2.6) \quad U^A(\psi) = \vec{\Gamma}(\lambda) \vec{w} - V(\lambda).$$

We follow [Grossman and Hart \(1983\)](#) and frame the problem in terms of finding a contract $\psi \in \Psi$ that implements an agreed-upon effort obligation λ at the lowest cost:

$$(2.7) \quad C(\psi) = \vec{\Gamma}(\lambda) \vec{c} = \vec{\Gamma}(\lambda) [\vec{w} + \vec{\delta}].$$

We ensure the existence of a non-degenerate optimal contract by requiring that all states occur with strictly positive probability under λ :

Definition 1. The effort obligation λ satisfies the *Full Support Condition (FSC)* at λ if $\Gamma_{ts}(\lambda) > 0, \forall ts \in S^2$.

In what is to follow, we focus only on the cost-minimizing aspect of the contracting problem. An optimal contract henceforth will thus refer to the cost-minimizing contract that implements a given effort λ .

For the Agent to accept a contract ψ , it must satisfy the Agent’s *Participation Constraint (PC)*:

$$(2.8) \quad U^A(\psi) = \vec{\Gamma}(\lambda) \vec{w} - V(\lambda) \geq U^0.$$

Since the Agent has unlimited liability, the PC always binds for the optimal contract. Combining (2.6) and (2.7), the Principal’s total expected cost is given by:

$$(2.9) \quad C(\psi) = V(\lambda) + \vec{\Gamma}(\lambda) \vec{\delta} + U^0.$$

Namely, costs consist of three components: the cost of getting the Agent's participation (U^0), the cost of effort ($V(\lambda)$), and the cost of ensuring parties are truthful ($\vec{\Gamma}(\lambda)\vec{\delta}$).

3. THE OPTIMAL SALES CONTRACT

This section outlines the constraints for good-faith and guile-free contracts. A general existence result for the optimal sales contract is provided.

3.1. Principal's Truthful-Reporting Constraints and Sales Contracts. Under a sales relationship information revelation arrangement outlined in steps (5) and (6) of the contracting game, since the Principal has observed the Agent's report on s before making his report, he will always report the t that minimizes his cost. Hence the *Principal's truthful-reporting constraint (PTR)* requires that:

$$(3.1) \quad c_{ts} \leq c_{t's}, \quad \forall t, t', s \in S.$$

This immediately implies:

Lemma 2. *Under a sales contract, the Principal's truthful-reporting constraint (3.1) implies that $c_{t's} = c_{ts} \forall t, t', s \in S$.*

Thus the Principal's cost c_{ts} can vary with s but not with t . Without loss of generality, we let p_s^S be the price that the Agent charges as a function of her signal (the superscript S denotes "sales contract") and hence $c_{ts} = p_s^S \forall t, s \in S$. Such an outcome corresponds to a sales contract whereby the Agent supplies a good or service to the Principal and informs him about the expected quality s . The Principal then pays a price p_s^S that varies only with the Agent's evaluation. The Agent's incentives to also report truthfully is described later.

We can directly integrate the PTR constraint into the definition of a sales contract by replacing the $n^2 \times 1$ cost vector \vec{c} with the $n \times 1$ cost-price vector \vec{p}^S , and define the set of sales contracts as:

$$(3.2) \quad \psi^S = \{\lambda, \vec{p}^S, \vec{w}\} \in \Psi^S \equiv [0, 1] \times \mathbb{R}^n \times \mathbb{R}^{n^2},$$

under the restriction of RBC (2.2). The conflict imposed in equilibrium is then $\delta_{ts}^S = p_s^S - w_{ts} \forall t, s \in S$.

3.2. On-path Constraints and Good-Faith Sales Contracts. Under \vec{w} , anticipating that she will report her signals truthfully on the equilibrium path, the Agent chooses an effort level to maximize her expected payoff in (2.6). Since $V(\cdot)$ is strictly convex, her optimal effort choice is determined by the first-order condition and the associated

complementary slackness condition:

$$\begin{aligned}\vec{\Gamma}\vec{w} - V'(\lambda) &\leq 0, \\ \lambda \left(\vec{\Gamma}\vec{w} - V'(\lambda) \right) &= 0.\end{aligned}$$

To implement a strictly positive effort λ , \vec{w} must thus satisfy the *incentive constraint for effort (ICE)*:¹⁵

$$(3.3) \quad \vec{\Gamma}\vec{w} = V'(\lambda).$$

To ensure that the Agent truthfully reports her signal s , the Principal must be able to punish the Agent if he believes that he has been short-changed. The punishment is inflicted through the conflict term $\vec{\delta}$. Without loss of generality (Lemma 1), we let the Principal be the party imposing all the conflict under a sales contract and denote the conflict by δ_{ts}^S – namely the Principal imposes a cost δ_{ts}^S on the Agent when he observes t after the Agent has reported s .

The Agent's on-path truthful-reporting constraint requires that after exerting effort λ and observing a signal $s \in S$, she cannot increase her expected wages by reporting a signal $s' \neq s$. Hence the Agent's *truthful-reporting constraint (ATR)* under the effort obligation λ is:

$$(3.4) \quad \left(p_s^S - \sum_{t \in S} r_{ts}(\lambda) \delta_{ts}^S \right) \geq \left(p_{s'}^S - \sum_{t \in S} r_{ts}(\lambda) \delta_{ts}^S \right), \quad \forall s, s' \in S.$$

where

$$r_{ts}(\lambda) = Pr[t|s, \lambda] = \frac{\Gamma_{ts}(\lambda)}{\sum_{t' \in S} \Gamma_{t's}(\lambda)}$$

is the conditional probability of the Principal receiving signal t given that the Agent has put in effort λ and observed signal s .

Notice that if δ_{ts}^S is always zero, then ATR (3.4) can only be satisfied if $p_s^S = p_{s'}^S, \forall s, s'$. When this happens, the Agent's wage is the same under any state and she will thus always put in zero effort (i.e. ICE (3.3) cannot be satisfied for $\lambda > 0$).¹⁶ Hence, conflicts are necessary (i.e. $\delta_{ts}^S > 0$ for some $ts \in S^2$) to induce positive effort by the Agent.

¹⁵One could have an inequality when $\lambda = 1$ by providing overly strong incentives. However, as we shall see, that only increases costs, so without loss of generality we suppose that ICE is always satisfied with equality.

¹⁶Note that $\sum_{ts \in S^2} \hat{\Gamma}_{ts} = \sum_{ts \in S^2} (\Gamma_{ts}^H - \Gamma_{ts}^L) = 1 - 1 = 0$; hence, for any vector $\vec{\alpha}$ in which every entry in the vector is the constant α , we have $\vec{\Gamma}\vec{\alpha} = 0 < V'(\lambda)$.

The constraints we have described thus far suppose that at every stage, there is no gain from any one-shot deviation under the hypothesis that both parties adhere to the contractual obligations (past and future). This is consistent with good-faith behaviors by self-interested parties. More precisely, we say that the self-interested parties are *acting in good faith* if their behaviors are as follows:

- (1) Under the hypothesis that both parties will be truthful in their signal reporting in steps (5) and (6) of the contracting game, it is in the Agent’s best interest in steps (2) and (3) to accept the contract and choose the contractual effort obligation.
- (2) Under the hypothesis that the Principal will be truthful in step (6) and that the Agent has exerted the contractual effort obligation in step (3), it is in the Agent’s best interest in step (5) to be truthful.
- (3) Under the hypothesis that the Agent acts in *good faith*, as given by 1 and 2 above, it is in the Principal’s best interest to be truthful in step (6).

These requirements are weaker than the requirement for a perfect Bayesian equilibrium because we are not allowing the Agent to jointly deviate from the contractual effort obligation and truthful-reporting.¹⁷ Instead, we only require that the contract satisfies the “no gain from one-shot deviation” property.¹⁸ In the context of a repeated game, such a property on the strategies is both necessary and sufficient for a subgame perfect Nash equilibrium (Abreu (1988)).

These requirements are consistent with empirical evidence on how individuals respond to immediate incentive but may not contemplate more complex deviations. For example, Vigna and Malmendier (2006) shows how individuals purchase long-term gym memberships believing that this will lead to more frequent gym visits; however many succumbed to short-run incentive to not visit the gym, which in turn implies that the pay-as-you-go membership plan would have been the more sensible one. More generally, laboratory experiments have shown that subjects have difficulty making state-contingent reasoning about hypothetical scenarios (Charness and Levin, 2009; Esponda and Vespa, 2014; Ngangoue and Weizsacker, 2015).

These papers illustrate how individuals do not always hold correct beliefs regarding their future actions. We use the same idea here but in a more positive fashion – that is, the good-faith Agent believes she will always be truthful in her future evaluations. This notion of good faith is a well-established legal doctrine requiring parties to act as if all involved

¹⁷A strategy of the Agent here consists of both an effort decision and the mapping from received signal to reported message.

¹⁸If one supposes different Agents, but with the same preferences, act at each stage then the good faith contract would form a perfect Bayesian equilibrium.

parties adhere to the contractual terms. At the same time, it also captures Williamson's notion of self-interest seeking by assuming that parties succumb to short-run incentives (as in [Vigna and Malmendier \(2006\)](#)), which in turn implies that contracts have to be designed so that there are no gains from any one-shot deviation.

Formally, we define good-faith contracts as the following:

Definition 2. The set of *good-faith sales contracts* that implements effort λ is the set of sales contracts $\Psi^{GF}(\lambda) \subset \Psi^S$ such that every $\psi^S \in \Psi^{GF}(\lambda)$ satisfies PC (2.8), ICE (3.3) and ATR (3.4) at λ . An *optimal good-faith sales contract* implementing effort λ , denoted by $\psi_\lambda^{GF} \in \Psi^{GF}(\lambda)$, satisfies:

$$C^{GF}(\lambda) \equiv C(\psi_\lambda^{GF}) = \min_{\psi^S \in \Psi^{GF}(\lambda)} C(\psi^S),$$

with the convention that $C^{GF}(\lambda) = -\infty$ if there is no feasible good-faith sales contract implementing λ ($\Psi^{GF}(\lambda) = \emptyset$).

Since the constraints are linear inequalities, $\Psi^{GF}(\lambda)$ is a convex and compact set if it is not empty; it immediately follows that an optimal good-faith sales contract exists if $\Psi^{GF}(\lambda)$ is non-empty. Since the Principal can provide effort incentives with the threat of conflict δ_{ts}^S if he has information regarding performance, we have:

Proposition 1. *There exists an optimal good-faith sales contract implementing effort λ , $\psi_\lambda^{GF} \in \Psi^{GF}(\lambda)$, if the Principal's signal provides information regarding the Agent's effort: $\hat{\gamma}_{\bar{t}} = \sum_{s \in S} \hat{\Gamma}_{\bar{t}s} \neq 0$ for some $\bar{t} \in S$.*

Proof. As noted above, it is sufficient to show that the set $\Psi^{GF}(\lambda)$ is not empty. Suppose $\hat{\gamma}_{\bar{t}} < 0$ (observing that $\sum_{t \in S} \hat{\gamma}_t = 0$ allows one to use the same argument for $\hat{\gamma}_t > 0$). Let $\delta_{ts}^S = \bar{\delta}$, $s \in S$ where $\bar{\delta} = V'(\lambda) / \hat{\gamma}_{\bar{t}}$, and set $\delta_{t's'} = 0$ for $t' \neq t, s, t' \in S$. Finally, set $p_s^S = \bar{p}$, $s \in S$ to ensure the PC is binding: $\bar{p} - \sum_{t,s \in S} \delta_{ts}^S \Gamma_{ts}(\lambda) - V(\lambda) = U^0$. By construction this contract satisfies the PC, ICE and ATR, and hence $\Psi^{GF}(\lambda) \neq \emptyset$. \square

3.2.1. Example: Good-faith Sales Contract with an Uninformed Principal. Proposition 1 shows that, unsurprisingly, when the Principal has information regarding the Agent's performance, the set of feasible good-faith sales contracts is non-empty and an optimal good-faith sales contract exists. What is less clear is whether it is still possible to implement a good-faith sales contract if the Principal has no information regarding performance?

We provide an example now that shows that this is possible. Let $S = \{0, 1\}$, effort $\lambda \in (\frac{1}{2}, 1)$ and consider the following information structure:

- (1) Nature draw $t = 0$ or 1 with equal probability.

- (2) $\Pr [s = t|t, o = H] = \pi > \frac{1}{2} \forall s, t \in \{0, 1\}$.
 (3) $\Pr [s = t|t, o = L] = 1 - \pi < \frac{1}{2} \forall s, t \in \{0, 1\}$.

Representing the probabilities in a matrix with the following convention:

$$\Gamma^o = \begin{bmatrix} \Gamma_{00}^o & \Gamma_{01}^o \\ \Gamma_{10}^o & \Gamma_{11}^o \end{bmatrix},$$

one can verify that the associated probability matrices are:

$$\Gamma^L = \begin{bmatrix} (1 - \pi)/2, & \pi/2 \\ \pi/2, & (1 - \pi)/2 \end{bmatrix},$$

$$\Gamma^H = \begin{bmatrix} \pi/2, & (1 - \pi)/2 \\ (1 - \pi)/2, & \pi/2 \end{bmatrix},$$

with

$$\hat{\Gamma} = (2\pi - 1)/2 \begin{bmatrix} 1, & -1 \\ -1, & 1 \end{bmatrix}.$$

In this example, the Principal by construction has no information regarding performance – the signal $t = 0$ and $t = 1$ occur with equal probability, and hence the Principal information cannot be directly used to set compensation. In the absence of information from t , the Agent is in a similar situation – she does not know how to interpret s , and hence the price term p_s^S cannot be used to provide incentives. Without loss of generality we can set the price term to a constant (to be determined): $p_s^S = \bar{p} \forall s$.

Instead, incentives are provided by conflicts – the Principal is truthful because of the assumption that conflicts are imposed costlessly. Since the problem is symmetric, we can set $\delta_{00}^S = \delta_{11}^S = \bar{\delta}$ and $\delta_{01}^S = \delta_{10}^S = \tilde{\delta}$. Let $\{\bar{p}, \bar{\delta}, \tilde{\delta}\}$ denote the contract with these terms. The Agent's expected payoff in equilibrium is:

$$U^A \left(\{\bar{p}, \bar{\delta}, \tilde{\delta}\} \right) = \bar{p} - \pi \bar{\delta} - (1 - \pi) \tilde{\delta} + \lambda (2\pi - 1) (\bar{\delta} - \tilde{\delta}) - V(\lambda).$$

ICE (3.3) is:

$$(3.5) \quad (2\pi - 1) (\bar{\delta} - \tilde{\delta}) = V'(\lambda).$$

(3.5) illustrates how the incentive to exert effort is created by imposing more conflict when the Principal disagrees with the Agent.¹⁹

¹⁹If $\pi < \frac{1}{2}$, then the opposite happens – effort incentives is created by imposing more conflict when the Principal agrees with the Agent.

Next, the Agent's truthful-reporting constraint ATR (3.4) for $s = 1$ is:²⁰

$$(3.6) \quad \begin{aligned} \bar{p} - \pi\bar{\delta} - (1 - \pi)\tilde{\delta} + \lambda(2\pi - 1)(\bar{\delta} - \tilde{\delta}) &\geq \bar{p} - \pi\tilde{\delta} - (1 - \pi)\bar{\delta} + \lambda(2\pi - 1)(\tilde{\delta} - \bar{\delta}) \\ &\iff (2\lambda - 1)(2\pi - 1)(\bar{\delta} - \tilde{\delta}) \geq 0 \end{aligned}$$

One can verify that the corresponding constraint for $s = 0$ is the same as (3.6). Since $\lambda > \frac{1}{2}$, (3.6) is subsumed by (3.5).

An optimal contract minimizes total conflict costs (see (2.9)) among the feasible contracts. From (3.5), this implies having

$$\begin{aligned} \delta_{00}^{GF} = \delta_{11}^{GF} = \tilde{\delta}^{GF} &= 0 \\ \delta_{01}^{GF} = \delta_{10}^{GF} = \bar{\delta}^{GF} &= \frac{V'(\lambda)}{2\pi - 1} \end{aligned}$$

\bar{p} is then set to ensure the participation constraint (2.8) binds:

$$\begin{aligned} \bar{p}^{GF} &= U^0 + V(\lambda) + (\pi - \lambda(2\pi - 1))\bar{\delta}^{GF}. \\ &= U^0 + V(\lambda) + V'(\lambda) \left(\frac{\pi}{(2\pi - 1)} - \lambda \right). \end{aligned}$$

We have thus solved for an optimal good-faith sales contract that implements λ . There are two lessons here. First, it is possible to design contract that provide incentives, even when the Principal has no information. Second, even though the setup seems to be a simple information revelation problem, it is sufficiently rich to capture quite complex and varied situations.

The example here is similar to Rahman's (2012) use of a trick question. Rahman considers the problem of how to ensure a monitor exerts effort in monitoring. In his example, the realization of t corresponds to the Principal secretly taking money from the till. He wants to make sure that the Agent in fact counted the money – if she did not, then she would not know if the Principal had changed the amount.²¹

3.3. Opportunism. Good-faith contracts have the feature that there are no gains from one-shot deviation, a feature we have identified with Williamson's notion of "self-interest".

²⁰We have multiplied the equation throughout by $\Pr[s = 1|\lambda] = \frac{1}{2}$.

²¹Rahman's mechanism is not quite right because it assumes that things tally with probability 1, but the whole point of doing a tally is that sometimes there is a discrepancy. However, as long as the distribution of discrepancies is known, then one can get correct probabilistic incentives from that mechanism.

When [Williamson \(1975\)](#) refers to the combination of self-interest with guile as opportunism, he explicitly intends for the notion of *guile* to refer to a more devious type of behavior.²² We capture guile in our model as the Agent combining deviation from the effort obligation with misrepresentation about her private information. This is aptly illustrated by the recent Volkswagen scandal – the car manufacturer first shirks from its obligation to develop “green” diesel automobile (deviation from effort obligation) and then implanted a software to cheat environmental testing systems and engaged in false advertising about their automobiles (misrepresentation about private information).²³

In this section, we develop the constraints necessary for a contract to be guile-free. A contract that is both guile-free and satisfies the good-faith constraints is robust against opportunism. Further notations are needed to denote the Agent’s reporting strategy in order to model guile as a strategic choice. Notice that the Agent’s reporting strategy is a mapping from a signal s to a report \hat{s} , and this can be represented by a $n^2 \times n^2$ linear transformation matrix Π on the contractual wage vector \vec{w} .²⁴ For example, suppose that $S = \{0, 1, 2\}$. The wage vector \vec{w} is thus:

$$\vec{w} = \left[\underbrace{w_{00}, w_{10}, w_{20}}_{s=0}, \underbrace{w_{01}, w_{11}, w_{21}}_{s=1}, \underbrace{w_{02}, w_{12}, w_{22}}_{s=2} \right]^T,$$

where the superscript T denotes the transpose. Consider a reporting strategy that stipulates the Agent to report $\hat{s} = 1$ when she observes $s = 0$ or $s = 1$, and to report $\hat{s} = 0$ when she observes $s = 2$. Reporting strategy Π will then be associated to the wage vector \vec{w}' , where:

$$\vec{w}' = \left[\underbrace{w_{01}, w_{11}, w_{21}}_{s=0 ; \hat{s}=1}, \underbrace{w_{01}, w_{11}, w_{21}}_{s=1 ; \hat{s}=1}, \underbrace{w_{00}, w_{10}, w_{20}}_{s=2 ; \hat{s}=0} \right]^T,$$

and the Agent’s expected wage with this reporting strategy under probability vector $\vec{\Gamma}(\lambda)$ is $\vec{\Gamma}(\lambda) \vec{w}'$, the inner product between $\vec{\Gamma}(\lambda)$ and \vec{w}' . It is readily observed that there exists a $3^2 \times 3^2$ matrix such that $\Pi \vec{w} = \vec{w}'$; the linear transformation matrix Π thus represents

²²[Williamson \(1975\)](#), page 26 quotes [Goffman \(1969\)](#), page 88 who states that opportunism involves “false or empty, that is, self-disbelieved, threats and promises” in the expectation that individual advantage will be realized.

²³See http://en.wikipedia.org/wiki/Volkswagen_emissions_scandal .

²⁴Formally, a reporting strategy of the Agent can be viewed by a $n \times n$ matrix π whose entries are either 1 or 0, each row has exactly one entry of 1, and $\pi_{s\hat{s}} = 1$ means that the Agent reports state \hat{s} when the true state is s ; the truthful-reporting strategy is represented by the identity matrix I . The transformation from π to Π is then as follows: for any $s, s' \in S$, $\Pi_{\mathcal{I}(ts), \mathcal{I}(ts')} = 1 \forall t \in S$ if $\pi_{ss'} = 1$, and 0 otherwise, where $\mathcal{I}(\cdot)$ is defined in fn. 9.

the aforementioned reporting strategy. Since $I\vec{w} = \vec{w}$, where I is the identity matrix, the truthful-reporting strategy is represented by $\Pi = I$.

Let Z be the set of possible reporting strategies Π for the Agent. The set Z has a cardinality of n^n , and hence the number of reporting strategies grows exponentially with the number of signals. Guile, the joint-deviation described above, involves a deviation in effort obligation to an effort vector $\lambda^g \neq \lambda$ together with a non-truthful reporting strategy $\Pi \in Z$. Hence, for a sales contract $\psi^S = \{\lambda, \vec{p}^S, \vec{w}\}$, the *guile-free constraint* is:

$$(3.7) \quad \vec{\Gamma}(\lambda)\vec{w} - V(\lambda) \geq \vec{\Gamma}(\lambda^g)\Pi\vec{w} - V(\lambda^g), \quad \forall \lambda^g \in [0, 1], \forall \Pi \in Z.$$

The left-hand side of (3.7) is the Agent's ex-ante expected payoff of adhering to the contract effort obligation and then reporting truthfully; the right-hand side is her expected payoff under effort λ^g and then using reporting strategy Π .

Notice that when λ^g in (3.7) is the contractual effort obligation, λ , we have ATR (3.4). On the other hand, when Π is the identity matrix (but allowing λ^g to vary now), (3.7) implies ICE (3.3). Hence the guile-free constraint in (3.7) subsumes the Agent's truthful-reporting constraint ATR and her incentive-compatibility constraint for effort obligation ICE. Thus, the guile-free constraint fully characterizes the potential for opportunistic behavior by the Agent (we do not use the term opportunism-free since it seems so awkward!).

Definition 3. The set of *guile-free sales contracts* that implements effort λ is the set of contract $\Psi^S(\lambda) \subset \Psi^S$ such that every $\psi^S \in \Psi^S(\lambda)$ satisfies PC (2.8) and the guile-free constraint (3.7). An *optimal sales contract* implementing effort λ , denoted by $\psi_\lambda^S \in \Psi^S(\lambda)$, satisfies:²⁵

$$C^S(\lambda) \equiv C(\psi_\lambda^S) = \min_{\psi^S \in \Psi^S(\lambda)} C(\psi^S),$$

with the convention that $C^S(\lambda) = \infty$ if there is no feasible guile-free sales contract implementing λ ($\Psi^S(\lambda) = \emptyset$).

An optimal sales contract is guile-free by construction, while a good-faith sales contract might not be. Hence, it follows that if the optimal good-faith contracts are not guile-free, then the need to ensure a contract is also guile-free increases the Principal's total expected cost.

Next, we show how the guile-free constraint can be simplified to provide necessary and sufficient conditions for the existence of an optimal sales contract. The representation in

²⁵It is readily observed that for a fixed effort obligation λ , the contract choice terms \vec{p}^S , \vec{w} and $\vec{\delta}$ in the optimal contracting program are linear in the objective, and also in the constraints PC and (3.7). Hence the problem is convex and there will be no gain from offering stochastic terms in the contracts.

(3.7) implies the guile-free constraint is an infinite set of inequalities due to the need for (3.7) to be satisfied for all $\lambda^g \in [0, 1]$. This can be reduced to a finite set of inequalities by observing that the Agent's optimal effort choice conditional upon her reporting strategy is unique. Since the number of reporting strategies is finite, we only need to check a finite number of inequalities.

Define the function $g : \mathfrak{R} \rightarrow \mathfrak{R}$ to be the maximum payoff of the Agent given the marginal return to effort x :

$$(3.8) \quad g(x) = \max_{\lambda \in [0,1]} \{\lambda x - V(\lambda)\}.$$

It follows immediately that $\forall \lambda^g \in [0, 1]$:

$$\begin{aligned} \vec{\Gamma}^L \Pi \vec{w} + g\left(\vec{\Gamma} \Pi \vec{w}\right) &\geq \vec{\Gamma}^L \Pi \vec{w} + \lambda^g \vec{\Gamma} \Pi \vec{w} - V(\lambda^g) \\ &= \vec{\Gamma}(\lambda^g) \Pi \vec{w} - V(\lambda^g). \end{aligned}$$

Under a wage vector $\Pi \vec{w}$, the Agent's *marginal effort incentive* is $\vec{\Gamma} \Pi \vec{w}$, and $\vec{\Gamma}^L \Pi \vec{w} + g\left(\vec{\Gamma} \Pi \vec{w}\right)$ gives the Agent's expected payoff under the optimal effort level given effort incentives $\vec{\Gamma} \Pi \vec{w}$. Thus we have:

Proposition 2. *The guile-free constraint in (3.7) consists of a finite set of inequalities. A sales contract $\psi^S = \{\lambda, \vec{p}^S, \vec{w}\} \in \Psi^S$ satisfies (3.7) for effort obligation λ if and only if:*

$$(3.9) \quad \vec{\Gamma}(\lambda) \vec{w} - V(\lambda) \geq \vec{\Gamma}^L \Pi \vec{w} + g\left(\vec{\Gamma} \Pi \vec{w}\right) \quad \forall \Pi \in Z.$$

$g(x)$ is strictly increasing and strictly convex when $x \in (0, V'(1))$; $g(x) = 0$ when $x \leq 0$, and $g(x) = x - V(1)$ if $x \geq V'(1)$. ICE (3.3) and ATR (3.4) are also satisfied when (3.9) is satisfied.

Let $\mathbb{G}(\lambda) \subset \mathfrak{R}^{n^2}$ be the set of wage vectors that satisfies condition (3.9). The following provides the existence condition for an optimal sales contract; it follows immediately from noting that $\mathbb{G}(\lambda)$ is a compact and convex set when it is non-empty.

Proposition 3. *Suppose that the full support condition FSC holds at λ . There exists an optimal sales contract that implements λ if and only if the set $\mathbb{G}(\lambda)$ is non-empty.*

A simple sufficient condition for the existence of a guile-free sales contract follows from the proof of Proposition 1. The contract used there satisfies ICE and has the feature that regardless of effort, the Agent's payoff is independent of her reporting strategy which thus satisfies the guile-free constraint trivially. Hence we have:

Corollary 1. *There exists an optimal sales contract implementing effort λ if the Principal's signal provides information regarding Agent's effort: $\hat{\gamma}_{\bar{t}} = \sum_{s \in S} \hat{\Gamma}_{\bar{t}s} \neq 0$ for some $\bar{t} \in S$.*

3.3.1. *Example: Guile-free Sales Contract with an Uninformed Principal.* As in Section 3.2.1, we ask the same question of whether it is also possible to implement a guile-free sales contract if the Principal only has information on the Agent's information. For the example in Section 3.2.1, suppose the Agent exerts zero effort *and* then always misrepresent her information (i.e reports $\hat{s} = 1$ when $s = 0$, and $\hat{s} = 0$ when $s = 1$). Her expected payoff is then $\bar{p}^* - (1 - \pi)\bar{\delta}^*$. The difference between this and her payoff under the optimal good-faith sales contract is:

$$\begin{aligned} U^A \left(\left\{ \bar{p}^{GF}, \bar{\delta}^{GF}, \tilde{\delta}^{GF} \right\} \right) - \left(\bar{p}^{GF} - (1 - \pi)\bar{\delta}^{GF} \right) &= (\lambda - 1)(2\pi - 1)\bar{\delta}^{GF} - V(\lambda) \\ &= (\lambda - 1)V'(\lambda) - V(\lambda) \\ &< 0, \end{aligned}$$

and hence the Agent benefits from opportunism. We can also show that $\mathbb{G}(\lambda)$ is empty because a single conflict level is doing a dual duty of inducing effort and ensuring truthful reporting.

However, implementation is possible under a similar environment if we expand the signal space. Consider the case with $S = \{0, 1, 2\}$, effort $\lambda \in \left(\frac{1}{2}, 1\right)$ and an information structure given by:

- (1) Nature draws $t = 0, 1$ or 2 with equal probability.
- (2) $\Pr[s = t | t, H] = \pi \forall s, t \in \{0, 1, 2\}$.
- (3) $\Pr[s = t + 1 | t, H] = 1 - \pi \forall s, t \in \{0, 1, 2\}$, where $t + 1 = 0$ if $t = 2$.
- (4) $\Pr[s = t | t, L] = 1 - \pi \forall s, t \in \{0, 1, 2\}$.
- (5) $\Pr[s = t + 1 | t, L] = \pi \forall s, t \in \{0, 1, 2\}$, where $t + 1 = 0$ if $t = 2$.
- (6) $\Pr[s = t + 2 | t, o] = 0 \forall s, t \in \{0, 1, 2\}$, $o \in \{L, H\}$, where $t + 2 = 0$ if $t = 1$, and $t + 2 = 1$ if $t = 2$.

We assume that $\pi > \frac{1}{2}$. The Principal's signal contains no information regarding the Agent's performance other than on how to infer the performance from the Agent's signal. The information structure implies that if $o = H$, the Agent is most likely going to observe the same signal as the Principal; if $o = L$, the Agent will most likely observe a slightly higher signal (i.e. $s = t + 1$); but regardless of performance, the Agent will never observe a much higher signal (i.e. $s = t + 2$) than the Principal.

Since $\Pr[s = t + 2] = 0$, the contract can specify an arbitrarily large value of conflict at these states (i.e. δ_{02}^S , δ_{10}^S and δ_{21}^S) without affecting the total expected cost of conflict. In

turn, since there is now a possibility of an arbitrarily high conflict cost for mis-reporting under any s , the Agent’s truthful-reporting constraints are trivially satisfied *regardless* of the effort she has previously exerted; hence the guile-free constraint is also trivially satisfied.

Similar to Section 3.2.1, the effort incentives are then provided by conflicts at states where $s = t + 1$, with:

$$\delta_{01}^S = \delta_{12}^S = \delta_{20}^S = \frac{V'(\lambda)}{(2\pi - 1)},$$

while optimally setting conflicts at states with $s = t$ (i.e. δ_{00}^S , δ_{11}^S and δ_{22}^S) to be zero. The payment \bar{p} is then set to satisfy the participation constraint with equality. This shows that $\mathbb{G}(\lambda)$ is not empty. There are many optimal contracts because the full support assumption is not satisfied. If one adds a small amount of noise in all states, then $\mathbb{G}(\lambda)$ is also compact, and there is a generically unique optimal contract.

3.4. Complexity Considerations. One way to think about the effect of guile is how strongly individuals respond to short-run incentives. In our model, the Principal would like to reinforce myopic behaviors by the Agent because it leads to good-faith behaviors which improves the efficiency of the contracts.

Another way to think about guile is that this is an expensive activity. Time and effort may be required to figure out how to game a particular contract; in turn, time and effort is also required to deter the gaming. One way to “quantify” the cost of these activities is by simply counting the number of additional constraints that the guile-free sales contract has to satisfy beyond those already satisfied under the good-faith sales contracts. This is illustrated in Table 3.1.

TABLE 3.1. Number of Constraints for the Good-Faith Sales and the Optimal Sales Contracting Problems

Number of Signals	Good Faith Sales Contract	Optimal Sales Contract
2	4	5
5	22	3126
10	92	$10^{10} + 1$

For a signal space of size n , the good-faith sales contracting problem has a total of $n^2 - n + 2$ constraints to satisfy.²⁶ On the other hand, the guile-free sales contract has the

²⁶1 constraint for ICE (3.3), 1 for PC (2.8), and $n(n - 1)$ truthful-reporting constraints for ATR (3.4), making it a total of $n^2 - n + 2$.

guile-free and participation constraints to satisfy; by Proposition 2, this amounts to a total of $n^n + 1$ constraints. This implies that as n increases, the number of constraints for the optimal sales contracting problem increases exponentially faster; with just $n = 10$ signals, there are already 10 billion constraints.

Moreover, this has not taken into account that the contracting problem for the optimal good-faith sales contract is linear, while that for the optimal sales contract is non-linear because of the guile-free constraint represented in (3.9). When we think of devious behaviors, we have in mind an individual who expends effort to find ways to game a system rather than enhancing the job performance. Given the complexity of this activity, it is not surprising that Principals cannot always anticipate all the ways a person may game a contract. Moreover, contracts that initially work well may become prone to opportunism over time as Agents learn to game a particular contract. The next section explores in greater details the constraint implied by potential opportunism.

4. OPPORTUNISM IN SALES CONTRACTS

We begin with an example to illustrate how the problem of guile arises in good-faith sales contracts. We then provide general conditions under which optimal good-faith sales contracts are susceptible to guile.

4.1. An Example with Two Signals. Let $S = \{A, U\}$ where A denotes an acceptable performance while U denotes an unacceptable performance. For simplicity, effort is assumed to be binary ($\lambda \in \{0, 1\}$), with the cost of high effort given by $\bar{V} \equiv V(1)$. In this setting, the Agent is perfectly informed about the outcome since $\lambda = 1$ implies $o = H$; $\lambda = 0$ implies $o = L$. Hence the purpose of her signal is to provide information regarding the Principal's perception of performance.

Suppose that the Agent's signal structure is given by:

$$(4.1) \quad Pr[s = A|o = H] = Pr[s = U|o = L] = \rho > \frac{1}{2}.$$

The Principal's signal is positively correlated with the Agent's information:

$$\begin{aligned} Pr[t = s|s, o = H] &= q, \\ Pr[t = s|s, o = L] &= q - \epsilon, \end{aligned}$$

with $q, q - \epsilon > 1/2$. As ϵ increases, the distortion in signals after the Agent shirks (which results in $o = L$) increases. The probability of each state is illustrated in Table 4.1.

Since $\rho, q - \epsilon > 1/2$, this implies that observing $t = A$ is good news regarding performance. One could have a fixed price contract, with all incentives are supplied via the

TABLE 4.1. Probabilities of States

$\lambda = 1$		
	$s = A$	$s = U$
$t = A$	ρq	$(1 - \rho)(1 - q)$
$t = U$	$\rho(1 - q)$	$(1 - \rho)q$

$\lambda = 0$		
	$s = A$	$s = U$
$t = A$	$(1 - \rho)(q - \epsilon)$	$\rho(1 - q + \epsilon)$
$t = U$	$(1 - \rho)(1 - q + \epsilon)$	$\rho(q - \epsilon)$

Principal imposing a cost when he observes $t = 0$. This is not in general optimal. In particular, if $q = 1$, then conflict costs can be greatly reduced by having the Agent charge a price as a function of s , and then the Principal punishes when his evaluation differs from the Agent's. Thus, even in the case in which the Agent perfectly observes performance, it may be more efficient condition upon a noisy signal s that is correlated with Principal's information.

Consider now the optimal good-faith sales contract. Since $s = A$ is more indicative of effort, the Agent would be paid a bonus when she reports A . To deter the Agent from always reporting A even when $s = U$, the Principal must punish the Agent when his signal suggests that the Agent is lying. This implies imposing a conflict δ at state $ts = UA$. The form of the optimal contract is summarized by Table 4.2.

TABLE 4.2. Optimal Good-Faith Sales Contract

Payoff of Agent		
	$s = A$	$s = U$
$t = A$	$p + b$	p
$t = U$	$p + b - \delta$	p

It is readily verified that the ATR (3.4) here is more stringent at $s = U$, which is represented by:

$$(4.2) \quad p \geq p + b - q\delta \quad \iff \quad \delta \geq \frac{b}{q}.$$

The level of conflict necessary to enforce truthful-reporting by the Agent is increasing in the bonus (b) and decreasing in the correlation in signals (q). Since conflict must be minimized for the optimal contract, (4.2) is set to hold with equality.

Next, the incentive for effort ICE (3.3) here can be written explicitly as:

$$(4.3) \quad p + \rho b - \rho(1 - q)\delta - \bar{V} \geq p + (1 - \rho)b - (1 - \rho)(1 - q + \epsilon)\delta$$

$$(4.4) \quad \iff b \geq \frac{q}{(2\rho - 1)(2q - 1) + \epsilon(1 - \rho)} \bar{V}.$$

The left-hand side of (4.3) is the Agent's expected payoff of adhering to the contractual terms, while the right-hand side is the payoff from exerting $\lambda = 0$ and then truthfully reporting her signal. (4.4) is obtained after substituting for $\delta = \frac{b}{q}$. The ICE must also bind for the optimal contract which then gives:

$$b^{GF} = \frac{q}{(2\rho - 1)(2q - 1) + \epsilon(1 - \rho)} \bar{V},$$

$$\delta^{GF} = \frac{1}{(2\rho - 1)(2q - 1) + \epsilon(1 - \rho)} \bar{V}.$$

Notice that increasing the information distortion after shirking (ϵ) increases the size of the bonus needed to induce effort, which in turn increases the level of conflict.

Finally, the fixed payment p does not affect any of the incentive constraints and is set to ensure the participation constraint PC binds:

$$p^{GF} + \rho b^{GF} - \rho(1 - q)\delta^{GF} - \bar{V} = U^0$$

$$\iff p^{GF} = U^0 + \bar{V} - \left(\frac{\rho(2q - 1)}{(2\rho - 1)(2q - 1) + \epsilon(1 - \rho)} \right) \bar{V}.$$

The Principal's expected cost of implementation is:

$$C^{GF} = U^0 + \bar{V} + \rho(1 - q)\delta^{GF}$$

$$= U^0 + \bar{V} + \frac{\rho(1 - q)}{(2\rho - 1)(2q - 1) + \epsilon(1 - \rho)} \bar{V}.$$

Let ψ^{GF} be the optimal good-faith contract which we have just derived. We consider if this contract is also guile-free. Let $u^A(\lambda, \Pi|\psi)$ be the Agent's expected payoff from exerting effort λ and using reporting strategy Π under a contract ψ . Let Π_A be the reporting strategy of always reporting $\hat{s} = A$ regardless of what is the underlying signal s ; and let I be the identity matrix that denotes the truthful-reporting strategy. Notice that:

$$u^A(0, \Pi_A|\psi^{GF}) = p^{GF} + b^{GF} - [(1 - \rho)(1 - q + \epsilon) + \rho(q - \epsilon)]\delta^{GF}$$

$$u^A(0, I|\psi^{GF}) = p^{GF} + (1 - \rho)b^{GF} - (1 - \rho)(1 - q + \epsilon)\delta^{GF}$$

Using $b^{GF} = q\delta^{GF}$, we have:

$$u^A(0, I|\psi^{GF}) = u^A(0, \Pi_A|\psi^{GF}) - \epsilon\rho\delta^{GF}.$$

Finally, using the fact that ICE binds ((4.3) holds with equality), we have $u^A(1, I|\psi^{GF}) = u^A(0, I|\psi^{GF})$, which in turn implies

$$u^A(1, I|\psi^{GF}) = u^A(0, \Pi_A|\psi^{GF}) - \epsilon\rho\delta^{GF}.$$

Hence, if $\epsilon > 0$, an opportunistic Agent can game the good-faith contract ψ^{GF} by shirking and then always reporting that the performance is acceptable. But such opportunism from guile will not be possible if $\epsilon = 0$. This turns out to be a more general feature which we will explore in Section 4.3.

The solution to deter the type of opportunism just described is to increase both b and δ until the contract ψ^* satisfies:

$$(4.5) \quad u^A(1, I|\psi^*) = U^0,$$

$$(4.6) \quad u^A(1, I|\psi^*) = u^A(0, I|\psi^*),$$

$$(4.7) \quad u^A(1, I|\psi^*) = u^A(0, \Pi_A|\psi^*).$$

From (4.6) and (4.7), we have $u^A(0, I|\psi^*) = u^A(0, \Pi_A|\psi^*)$ which gives:

$$\begin{aligned} p^* + (1 - \rho)b^* - (1 - \rho)(1 - q + \epsilon)\delta^* &= p^* + b^* - [(1 - \rho)(1 - q + \epsilon) + \rho(q - \epsilon)]\delta^* \\ \iff \delta^* &= \frac{b^*}{q - \epsilon}. \end{aligned}$$

Comparing with (4.2), when $\epsilon > 0$, the required δ is higher than that required to just satisfy the ATR. Substituting $\delta^* = \frac{b^*}{q - \epsilon}$ into (4.3) and setting it to equality as in (4.6), we get:

$$\delta^* = \frac{1}{(2\rho - 1)(2q - 1) + \epsilon(2 - 3\rho)} \bar{V}.$$

The Principal's expected cost of implementation is then:

$$\begin{aligned} C^* &= U^0 + \bar{V} + \rho(1 - q)\delta^* \\ &= U^0 + \bar{V} + \frac{\rho(1 - q)}{(2\rho - 1)(2q - 1) + \epsilon(2 - 3\rho)} \bar{V} > C^{GF}. \end{aligned}$$

The final inequality follows from the fact that $\rho > \frac{1}{2} \implies (1 - \rho) > (2 - 3\rho)$ since $\rho \in (1, 1/2)$. Thus, the introduction of the guile-free constraint increases the implementation cost if and only if $\epsilon > 0$.

4.2. Guile in Good-Faith Sales Contracts. In this section we explore the effect of opportunism more generally. Notice that given a sales contract $\psi^S = \{\lambda, \bar{p}^S, \bar{w}\}$, we can define the Agent's expected payoff from effort, $\lambda^g \in [0, 1]$, and reporting strategy, $\Pi \in Z$,

by:

$$u^A(\lambda^g, \Pi|\psi) = \text{wage}(\Pi|\psi) + \lambda^g \times \text{bonus}(\Pi|\psi) - V(\lambda^g),$$

where $\text{wage}(\Pi|\psi) = \vec{\Gamma}^L \Pi \vec{w}$ is the fixed component of compensation, while $\text{bonus}(\Pi|\psi) = \vec{\tilde{\Gamma}} \Pi \vec{w}$ is the component that varies with effort. This allows us to use a two-dimensional representation of how different reporting strategies, Π , affect compensation. Define:

$$WB(\psi) = \{(\text{wage}(\Pi|\psi), \text{bonus}(\Pi|\psi)) \mid \Pi \in Z\},$$

as the set of possible $(\text{wage}, \text{bonus})$ -pairs under \vec{w} that arise from different reporting strategies. Since the set of reporting strategies Z is finite, $WB(\psi)$ contains a finite number of points. The $(\text{wage}, \text{bonus})$ -pair associated to the truthful-reporting strategy, $\Pi = I$, is denoted by $\text{wage}^* = \vec{\Gamma}^L \vec{w}$ and $\text{bonus}^* = \vec{\tilde{\Gamma}} \vec{w}$.

We can illustrate the properties of feasible sales contracts in terms of a $(\text{wage}, \text{bonus})$ graph, as is shown in Figure 4.1. Consider first an optimal good-faith sales contract, ψ^{GF} , at which the Agent chooses effort λ and then truthfully reports her information, and she gets compensation $(\text{wage}^*, \text{bonus}^*)$, as illustrated in this figure. At such a contract:

$$\begin{aligned} \text{bonus}^* &= V'(\lambda), \\ u^A(\lambda, I|\psi^{GF}) &= U^0, \end{aligned}$$

where we have normalized U^0 to zero for the figure. Since ψ^{GF} satisfies ATR (3.4), any reporting strategy must give a payoff that is less than or equal to U^0 while holding effort fixed at λ . This corresponds to the requirement:

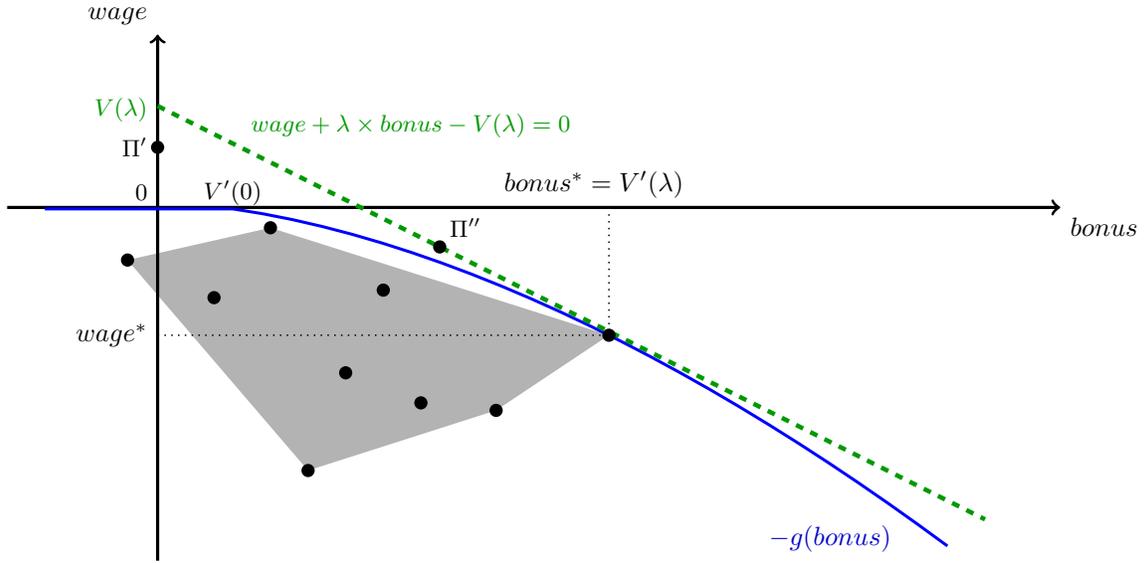
$$(4.8) \quad WB(\psi) \subset WB^{GF}(\psi) = \{(\text{wage}, \text{bonus}) \mid \text{wage} + \lambda \times \text{bonus} - V(\lambda) \leq U^0\}.$$

This is the requirement that all elements in $WB(\psi)$ be below the dashed line in Figure 4.1. Since $WB^{GF}(\psi)$ is a convex set, the convex hull of the finite number of points in $WB(\psi)$ is also in this set, as illustrated by the shaded region in the figure.

Guile corresponds to strategic manipulation via a “double deviation” that entails shirking, followed by misrepresentation of the Agent’s information. Information misrepresentation is equivalent to selecting a point in $(\text{wage}, \text{bonus}) \in WB(\psi)$ that is not $(\text{wage}^*, \text{bonus}^*)$. Using the function $g(\cdot)$ defined by (3.8), it is not optimal to carry out such strategic manipulation if and only if:

$$WB(\psi) \subset WB^*(\psi) = \{(\text{wage}, \text{bonus}) \mid \text{wage} + g(\text{bonus}) \leq U^0\}.$$

FIGURE 4.1. Guile in Good Faith Contracts



The set $WB^*(\psi)$ is given by the points $(wage, bonus)$ on or below the graph:

$$(4.9) \quad wage + g(bonus) = U^0,$$

The solution to (4.9) with $U^0 = 0$ is given by the curved solid graph in Figure 4.1.

As one can see, the set $WB^*(\psi)$ is a strict subset of $WB^{GF}(\psi)$, hence the set of guile-free sales contracts is smaller than the set of feasible good-faith sales contracts. In particular, there are a number of natural necessary conditions for a guile-free contract. First, suppose that there is a reporting strategy Π' (see Figure 4.1) that gives rise to a *wage* that is better than the outside option:

$$wage = \Gamma^L \Pi' \vec{w} > U^0,$$

then the Agent is better off choosing zero effort combined with the reporting strategy Π' , as illustrated in Figure 4.1.

Next, given that the payoff is linear in the contract terms, the ATR (3.4) must bind for some reporting strategy under the optimal good-faith sales contract. If this reporting strategy gives a $(wage, bonus)$ different from $(wage^*, bonus^*)$, as is the case for Π'' in Figure 4.1, then the Agent can benefit from opportunism via guile. This is because the Agent is indifferent between reporting strategy Π'' and truthful-reporting strategy (I) when holding effort fixed at λ . However, the convexity of the effort cost function implies that the Agent is

strictly better off reducing effort and reporting Π' . These observations imply the following necessary conditions for a contract to be guile-free:

Proposition 4. *An optimal good-faith sales contract $\psi^{GF} = \{\lambda, \bar{p}^{GF}, \bar{w}^{GF}\}$ is not guile-free if either of these conditions hold:*

- (1) *There is a reporting strategy, Π , resulting in a wage that is greater than the outside option: $\Gamma^L \Pi \bar{w}^{GF} > U^0$.*
- (2) *There is a reporting strategy, Π , giving the Agent a payoff equal to his outside option, U^0 , at the agreed upon effort λ , but resulting in a different wage: $\Gamma^L \Pi \bar{w}^{GF} \neq \Gamma^L \bar{w}^{GF}$.*

In the example in Section 4.1 above, both of these cases occur when $\epsilon > 0$. This in turn implies that the requirement that a contract be guile-free increases costs relative to just having an optimal good-faith sales contract.

4.3. Information Structure and Guile: The Expert-Agent. The presence of guile is *not* a mere artifact of combining moral hazard with asymmetric information in the form of subjective performance evaluation; it also depends on the underlying information structure and the contractual relationship. In this section, we provide a class of information structures which we call the *expert-Agent* environment in which guile does not exacerbate opportunism.

Definition 4. The information structure is an *expert-Agent* environment if for all outcomes $o \in \{L, H\}$, the probability of each state occurring can be represented by $\Gamma_{ts}^o = \beta_s^o r_{ts} \forall t, s \in S$, where $\beta_s^o = \sum_{t \in S} \Gamma_{ts}^o$, and $r_{ts} = Pr[t|s]$ is the conditional probability of the Principal observing signal t when the Agent observes s .

One can always write $\Gamma_{ts}^o = \beta_{ts}^o r_{ts}^o$ under any information structure. The expert-Agent environment restricts that the correlation r_{ts}^o is independent of the outcome o - intuitively, this condition is the requirement that the correlation of Agent and Principal's signals are independent of Agent performance, which can be the case where parties rely upon a common measurement system unaffected by the choice of effort, λ . Notice that in the example in Section 4.1, $\epsilon = 0$ corresponds to an expert-Agent environment.

Proposition 5. *In an expert-Agent environment, every good-faith sales contract is guile-free - that is, if a sales contract ψ^S satisfies ICE (3.3) and ATR (3.4), then ψ^S also satisfies the guile-free constraint (3.7).*

Proposition 5 illustrates that guile is not just due to the presence of asymmetric information. For example, suppose the Agent knows o perfectly, while the Principal only has a very noisy signal about what the Agent observes.²⁷ There is a high degree of asymmetric information about the outcome in this case – one party (the Agent) perfectly knows the outcome while the other party (the Principal) has almost no information about it. However, according to Proposition 5, the issue of guile turns out to be irrelevant here.

Instead, guile arises when the Agent has the ability to manipulate and create asymmetric information about the information flow to the players. The result in Proposition 5 rests on the observation that in the expert-Agent information structure, the Agent’s on-path truthful-reporting constraint ATR (3.4), is always independent of the Agent’s effort. This implies that the Agent is unable to create a divergence in information flow to the players in a way that is known only to the Agent but not to the Principal.

5. THE AUTHORITY CONTRACT

This section explores the implication of changing the order of information revelation in the contractual relationship on both the contract form and the possibility of opportunism. We consider the *authority relationship* where at step (5) in the contracting game described in Table 2.1, the Principal makes a report \hat{t} regarding his signal t first. Then at step (6), upon observing the Principal’s report \hat{t} , the Agent gives her report \hat{s} on her signal s . The rest of the contracting game remains unchanged.

Consider the Agent’s truthful-reporting constraint first. Since effort is sunk and the outcome has been realized, the Agent cares only about her net compensation. Thus the Agent’s truthful-reporting (ATR) constraint is:

$$(5.1) \quad w_{ts} \geq w_{ts'} \quad , \quad \forall t, s, s' \in S.$$

This immediately implies:

Lemma 3. *In an authority relationship in which the Principal reports his signal t first, the Agent’s truthful-reporting constraint ATR (5.1) requires that $w_{ts} = w_{ts'} \quad \forall t, s, s' \in S$.*

This implies that the Agent’s net compensation is completely determined by the Principal’s evaluation since w_{ts} can only vary with t . This contractual outcome resembles an authority contract that is typically observed in employment relationships in firms – the

²⁷Formally, let $S = \{0, 1\}$, and $\beta_0^L = \beta_1^H = 1$; hence $s = 0$ (respectively $s = 1$) is perfectly indicative of outcome L (respectively H). In addition, let $r_{00} = r_{11} = \frac{1}{2} + \varepsilon$ and $r_{01} = r_{10} = \frac{1}{2} - \varepsilon$ where $\varepsilon > 0$ is very small; hence the Principal has a very noisy signal about s . Note that ε cannot be 0; if not, there will be no contract that satisfy both ICE and ATR at the same time.

manager has the authority to evaluate the worker's performance and determine her compensations. Since the wage term w_{ts} varies with only t , we can define a wage-price term p_t^A that varies only with the Principal's report and let $w_{ts} = p_t^A \forall t, s \in S$.²⁸

Analogous to the sales contracts, we can directly integrate ATR (5.1) into the definition of an authority contract by replacing the $n^2 \times 1$ wage vector \vec{w} with the $n \times 1$ wage-price vector \vec{p}^A . The set of *authority contracts* is thus:

$$(5.2) \quad \psi^A = \left\{ \lambda, \vec{c}, \vec{p}^A \right\} \in \Psi^A \equiv [0, 1] \times \mathfrak{R}^{n^2} \times \mathfrak{R}^n,$$

with the restriction of RBC (2.2). The conflict in equilibrium is thus $\delta_{ts}^A = c_{ts} - p_t^A \geq 0, \forall t, s \in S$.

Notice that since the Agent's compensation is independent of her report in an authority contract, truthful-reporting is always a best response regardless of the effort that she has exerted. This implies the following result:

Proposition 6. *An authority contract satisfies the guile-free constraint in (3.7) if it satisfies the incentive constraint for effort ICE in (3.3).*

Next, to give incentive for the Principal to evaluate truthfully, the Agent has to be able to punish the Principal if she feels that her reward does not match up to her own evaluation of the performance. This is modeled as the Agent inflicting a conflict through δ in states in which her signal s indicates a likely misreporting by the Principal. Without loss of generality (Lemma 1), we let the Agent be the party who is imposing all the conflict under an authority contract. The Principal's truthful-reporting constraint (PTR) here requires that conditional upon the Agent exerting the effort obligation λ and upon the Principal observing a signal $t \in S$, he cannot reduce his costs by reporting a signal $t' \neq t$:

$$(5.3) \quad \left(p_t^A + \sum_{s \in S} q_{ts}(\lambda) \delta_{ts}^A \right) \leq \left(p_{t'}^A + \sum_{s \in S} q_{ts}(\lambda) \delta_{t's}^A \right), \quad \forall t, t' \in S,$$

where

$$q_{ts}(\lambda) = Pr[s|t, \lambda] = \frac{\Gamma_{ts}(\lambda)}{\sum_{s' \in S} \Gamma_{ts'}(\lambda)},$$

is the conditional probability of the Agent receiving signal s given that she has put in effort λ and the Principal has observed signal t .

Definition 5. The set of authority contracts that implements effort λ is the set of authority contracts $\Psi^A(\lambda) \subset \Psi^A$ such that every $\psi^A \in \Psi^A(\lambda)$ satisfies PC (2.8), ICE (3.3) and PTR

²⁸The superscript-“A” stands for “authority”.

(5.3). An *optimal authority contract* that implements effort λ , denoted by $\psi_\lambda^A \in \Psi^A(\lambda)$ satisfies:

$$C^A(\lambda) \equiv C(\psi_\lambda^A) = \min_{\psi^A \in \Psi^A(\lambda)} C(\psi^A),$$

with the convention that $C^A(\lambda) = -\infty$ if there is no feasible authority contract implementing λ ($\Psi^A(\lambda) = \emptyset$).

The participation constraint PC ensures that costs are bounded from below, and the fact that the set of constraints are linear implies that $\Psi^A(\lambda)$ is a convex set. Hence, the existence of an optimal authority contract implementing λ depends only upon whether or not the set $\Psi^A(\lambda)$ is non-empty. The following provides the condition for the existence of an optimal authority contract.

Proposition 7. *There exists an optimal authority contract that implements λ if and only if there exists some $t \in S$ such that $\sum_{s \in S} \hat{\Gamma}_{ts} \neq 0$.*

This result is intuitive. Under an authority contract, only the Principal can modulate the rewards that the Agent receives. Hence, he can provide performance incentives if and only if his signal is informative of effort, a condition that corresponds to $\sum_{s \in S} \hat{\Gamma}_{ts} \neq 0$. The Principal's truthful-reporting can always be ensured by a contract in which the Agent imposes a conflict cost that is equal to the value of the bonus (relative to the lowest payment) when the bonus is not paid. This makes the Principal indifferent among reporting any signal which then trivially satisfies PTR (5.3). This implies that the set $\Psi^A(\lambda)$ is always non-empty whenever $\sum_{s \in S} \hat{\Gamma}_{ts} \neq 0$. If there is some correlation between the players' signal, then the optimal contract exploits this correlation to reduce the expected conflict costs.

In standard agency theory, having higher quality information regarding the Agent's effort leads to more efficient contracts because the Principal can better provide effort incentives without subjecting the risk-averse Agent to too much risk. In our model, the Agent is risk-neutral, and the main tradeoff here lies between incentives for effort and the cost of conflicts needed to induce truthful-reporting by the Principal. The less informative is the Agent's information regarding the Principal's information, the more conflict is required in the relationship.

In section 3.3.1 we provided an example of a sales contract under which the Principal has no information, but there may exist a sales contract implementing positive effort. Proposition 7 implies that no authority contract exists in that case. Notice that the converse cannot hold, namely if there exists an authority contract implementing effort $\lambda > 0$, then there also exists a sales contract implement the same effort. This can be achieved with a fixed wage contract under which the Principal uses punishments to provide

the appropriate effort level. Since the Principal is indifferent over which costs are imposed, such a contract is incentive compatible:

Corollary 2. *If there exists an authority contract implementing effort λ , then there exists a guile-free sales contract implementing the same effort.*

These sales contracts in effect entail fixed wages, with worker facing a punishment in some states – this corresponds to the behavior under the well-known class of efficiency wage models that we now consider in more details.

5.1. Example: Efficiency Pay and Efficiency Wage Contracts. In Sections 3.2.1 and 3.3.1, we provided examples of implementation with a sales contract when the Principal is completely uninformed. From Proposition 7 we know that there does not exist an authority contract implementing positive effort. In this section, we consider the converse case where the Agent is completely uninformed ($\Gamma_{ts}^o = \Gamma_{ts'}^o, \forall t, s, s' \in S, o \in \{H, L\}$), but the conditions for the existence of an authority contract (Proposition 7) are satisfied. In contrast to the case of an informed Agent, in this case there always exists both authority and sales contracts that implement the desired effort. In this section we derive and compare both types of contracts and then explicitly link the sales contract to the canonical efficiency wage model of Shapiro and Stiglitz (1984).

Let:

$$\begin{aligned}\gamma_t^o &= \sum_{s \in S} \Gamma_{ts}, \\ \hat{\gamma}_t &= \gamma_t^H - \gamma_t^L,\end{aligned}$$

where the corresponding row vectors are $\vec{\gamma}^o = [\gamma_0^o, \dots, \gamma_{n-1}^o]$. For ease of exposition, we assume that $\gamma_t^o > 0$ and the signals are indexed by the likelihood ratio so that:

$$(5.4) \quad \frac{\gamma_{t+1}^H}{\gamma_{t+1}^L} > \frac{\gamma_t^H}{\gamma_t^L}, \quad \forall t = 0, \dots, n-2,$$

and $\gamma_t^H \neq \gamma_t^L \forall t \in S$. This implies that each signal is either good news ($\frac{\gamma_t^H}{\gamma_t^L} > 1$) or bad news ($\frac{\gamma_t^H}{\gamma_t^L} < 1$) regarding performance. Consider first an authority contract, where the Agent's wage is a function of the Principal's report t . Without loss, we can represent the wage as:

$$w_t = \bar{w} + b_t,$$

where $b_t \geq 0$ is a bonus payment at t .

Next, since the Agent's information is uninformative, the conflict that the Agent imposes on the Principal cannot depend on s – that is, $\delta_{ts}^A = \delta_{ts'}^A \forall t, s, s' \in S$. In turn, the Principal's

truthful-reporting constraint can only be satisfied if his costs are independent of his report, which means that the conflict terms must satisfy:

$$(5.5) \quad \delta_{ts}^A = \left\{ \max_{t' \in S} b_{t'} \right\} - b_t \quad \forall t, s \in S.$$

Let $\psi^{EP} = \{\lambda, \bar{w}, \vec{b}\} \in \Psi^{EP}(\lambda) = [0, 1] \times \mathfrak{R} \times \mathfrak{R}_+^n$ be an authority contract offered by the Principal, where $\vec{b} = [b_0, \dots, b_{n-1}]^T$, and the associated conflict costs are given by (5.5). Under ψ^{EP} , regardless of the Principal's report, his cost is always $\bar{w} + \bar{b}(\psi^{EP})$ where $\bar{b}(\psi^{EP}) = \max_{t' \in S} b_{t'}$.

The optimal efficiency pay contract that implements λ solves:

$$(5.6) \quad C^{EP}(\lambda) = \min_{\psi^{EP} \in \Psi^{EP}(\lambda)} \bar{w} + \bar{b}(\psi^{EP})$$

subject to:

$$\begin{aligned} \vec{\gamma} \vec{b} &= V'(\lambda), \\ \bar{w} + (\vec{\gamma}^L + \lambda \vec{\gamma}) \vec{b} - V(\lambda) &\geq U^0 \end{aligned}$$

where the first constraint is the effort incentive constraint ICE and the second constraint is the participation constraint PC.

Proposition 6 in MacLeod (2003) shows that the optimal contract here entails paying a bonus $b_t = \bar{b}(\psi^{EP})$ in all states except for the one that is most informative of a bad performance (i.e. $t = 0$). This result provides one explanation of why managers tend to rank most employees as “above average” and penalize only the worst performance. More formally, we have:

Proposition 8. *Suppose that (i) there exists some $t \in S$ such that $\sum_{s \in S} \hat{\Gamma}_{ts} \neq 0$ (i.e. Principal's signal is informative), (ii) $\Gamma_{ts}^o = \Gamma_{ts'}^o, \forall t, s, s' \in S, o \in \{H, L\}$ (i.e. Agent's signal is uninformative), and signals $t \in S$ are ordered by their likelihood ratios as in (5.4). Then there exists an optimal authority contract that implements $\lambda, \psi_\lambda^{EP}$, with the following form:*

- (1) *The Agent is paid a fixed wage \bar{w} in all states, and a bonus $b_t = \bar{b} \forall t > 0$.*
- (2) *The bonus satisfies:*

$$\bar{b} = \frac{V'(\lambda)}{-\hat{\gamma}_0}.$$

- (3) *The fixed wage satisfies:*

$$\bar{w} = U^0 + V(\lambda) - \left(1 - \gamma_0^L - \lambda \hat{\gamma}_0\right) \bar{b}.$$

(4) *Total cost is given by:*

$$C^{EP}(\lambda) = U^0 + V(\lambda) + (\gamma_0^L + \lambda\hat{\gamma}_0)\bar{b}.$$

Since $\sum_{t \in S} \hat{\gamma}_t = 0$, $\hat{\gamma}_0 < 0$ and hence $\bar{b} > 0$. Moreover, conflict occurs only at $t = 0$, the signal that is most indicative of poor performance. Thus the the expected social loss is $(\gamma_0^L + \lambda\hat{\gamma}_0)\bar{b}$, which in turn implies expression for the total cost.

Notice that from Lemma 1, the contract can also be implemented with a sales contract. The Principal pays a fixed wage $w^{EW} = \bar{w} + \bar{b}(\psi^{EP*})$ in every state and imposes a conflict cost $\delta_0 = \bar{b}$ on the Agent when $t = 0$. This is very similar to the “efficiency wage” contract of [Shapiro and Stiglitz \(1984\)](#), where the firm pays a fixed wage every period, and when the worker is found to shirk (i.e. $t = 0$ here), she is fired and replaced by an identical worker. [Shapiro and Stiglitz \(1984\)](#) take the probability of detecting cheating as fixed and work out the equilibrium level of unemployment consistent with ensuring high effort in equilibrium. In their model, unemployment plays exactly the same role as conflict plays in our model. This result is also consistent with recently experimental work by [Kaur \(2014\)](#) who provides evidence that workers can impose costs upon prospective employers if they believe they are treated unfairly.

6. CONCLUDING DISCUSSION

A challenge for agency theory is to explain not only the wide diversity in observed contracts, but also the many dysfunctional behaviors that have been documented over time. These observations are difficult to explain within the canonical principal-agent model because contract designers are assumed to anticipate strategic behaviors by Agents and would then design the contract to deter these behaviors accordingly. In this paper, we revisit the basic agency model with subjective evaluation to show that the order in which information is revealed by parties provides sharp predictions regarding the contract form, which in turn provides one possible explanation for why Principals may offer contracts that can be gamed.

The importance of timing upon the set of equilibria is highlighted in the work of [Myerson \(1986\)](#). He provides a general analysis of multi-stage games with communication and requires equilibria to be free from strategic manipulation – a combination of strategic information miss-representation combined with a change in actions relative to the agreed upon equilibrium. He shows that small changes in the game form can have a large impact upon the set of equilibria.

In this paper, we focus upon a game form that is specific to the agency problem which allows us to provide a concrete interpretation of strategic manipulation that corresponds precisely to Williamson’s notion of opportunism. In line with [Myerson \(1986\)](#), we find that changes in the timing of information release has a large impact upon the set of equilibria. Contracts that require the Principal to reveal his information before the Agent are called authority contracts. The resulting contract form is consistent with a firm’s right to manage under which it is free to set compensation for its employees as a function of its private information.

In principle, firms can elicit their employees’ information (for example, through the use of 360-degree evaluation systems), yet as [Williamson \(1991\)](#) observes, firms operate under forbearance law which makes it impossible for a firm to commit to any compensation system.²⁹ An employer is always free to change working conditions as a function of the information he holds. An interesting implication is that if the Agent has no private information, then the optimal authority contract entails punishing employees if and only if the worst signal of performance occurs, leading to what [MacLeod \(2003\)](#) calls the “Lake Wobegone effect” – the well-known tendency for managers to rate most employees as “above average”.³⁰

A feature of authority contracts is that they are also free from opportunism by the Agent – strategic manipulation via shirking accompanied with misrepresentation of private information is never optimal when an authority contract is free from gains through any one-shot deviation. The necessary and sufficient conditions for the implementation of an authority contract are the requirements that the Principal has some information regarding the Agent’s performance and that the Agent can also impose costs upon the Principal when she believes that she has been inadequately compensated. These costs can take a variety of forms, including a decrease in future effort (e.g. a morale effect, as in the case of the police unions discussed in [Mas \(2006\)](#)), or directly reducing the quality of output ([Mas, 2008](#)). Other possible avenues include via peer affects as in [Deb et al. \(2016\)](#), or through the use of stochastic contracts as in [Maskin and Tirole \(1999\)](#) and [Lang \(2016\)](#).

These results complement the work on relational contracts where the key issue is about how constraints on the size of the potential conflict affect the form and efficiency of contracts ([Levin, 2003](#); [Li and Matouschek, 2013](#)). In all of these cases, the conceptual framework is one in which the Principal sets compensation while the Agent responds if she feels unfairly

²⁹See pg 275.

³⁰See [Milkovich et al. \(2011\)](#).

treated. Such a contract form corresponds to an authority contract in our model where the Principal reveals his private information first.

However, if the order of information revelation is changed, a very different set of issues arises. Payment is now set by the Agent as opposed to the Principal and hence, we call these sales contracts. For example, the price of expert advice, such as in financial consulting and healthcare, is often determined based on the value set by the seller. We show that contracting via sales contracts leads to potential strategic manipulation corresponding to Williamson's notion of opportunism.

It is worth observing that what we have defined as opportunism differs from the behaviors modeled in the celebrated multi-tasking model of [Holmström and Milgrom \(1991\)](#). That model corresponds to what we have called an authority contract in which the Principal sets compensation as a function of the signals received. The difference now is that the Agent has many dimensions of effort, and the contract may not be optimal if the Principal is not aware of some of the other actions available. For example, rewarding a contractor for timely completion of a project can result in skipping on quality. In this case the Agent is not explicitly misrepresenting information, but merely reallocating time from some tasks to increase on time performance. In the online Appendix we extend the model to allow for a special case of the multi-tasking, namely malfeasance - actions that have no value, but affect the information flow to the Principal. We show that malfeasance can make opportunism more severe because it provides the Agent with more ways to manipulate the Principal's information.

The recent Volkswagen scandal provides an illustration of how opportunistic behavior and malfeasance work together. The car manufacturer Volkswagen breached its obligation to the Environmental Protection Agency (EPA) in the US by creating a software that cheats the EPA's testing system, while at the same time engaging in deceitful marketing of their diesel automobiles as "green". We can view the EPA's regulatory procedure as a "good faith" regime where they monitor an automobile's emissions level with a stationary testing unit. Such a system is less expensive than using road tests. Their existence provides an incentive for compliance - no gain from one-shot deviation in the model - but it does rely upon car manufacturers "to act in good faith" and to not devise more sophisticated methods to game the testing system, as Volkswagen did.

In the context of the multi-tasking model, Volkswagen's behavior can be viewed as the logical response to the incentives provided by the EPA. However, the testing system is not intended to be a reward/punishment system - passing the testing does not absolve VW's obligation to meet pollution standards, it is in addition to other systems used by the

EPA to encourage compliance with the law. The European response to the same issue is however consistent with the multi-tasking model. Rather than place an obligation of good faith performance, the European testing system is viewed as an absolute reward system which, in light of the behavior by the car companies, requires some “loop-hole” fixes to encourage compliance with pollution standards.³¹

Consider an even more pertinent example – what happens when an individual buys a good from a seller in “good faith”, only to later realize that he has bought a stolen good? Should the original owner be allowed to recover the good, and should the individual also be liable for legal penalties for the transaction? Interestingly, [Levmore \(1987\)](#) documents how the law’s treatment of these obligations varies across jurisdictions.

In the context of our model, the effort of the seller (Agent) here can be mapped to the choice between producing or stealing the good, and the information revelation is on how the seller obtained the good. The efficient solution here is an environment where sellers always act in good faith. Buyers then only have to incur the cost of inspecting the quality of the good before accepting the price demanded by the seller. It is clear that there would be significant inefficiencies if the buyers also have to ensure that the goods were not stolen in the first place.

There can be a variety of reasons to why individuals would act in good faith and expect others to also do so. In our model, we show that one reason this can occur is because the number of constraints to be checked to ensure that a contract is free from opportunism increases exponentially with the number of available signals regarding performance. This implies that there is a cost to such opportunistic behaviors, and thus it may be rational for the Principal to simply offer “good-faith contracts” if he believes that the Agent will act in good faith. Likewise, while we recognize that theft and resale is financially lucrative, we still reasonably expect sellers to act in good faith and not adopt this opportunistic strategy in some situations but not in others – think about purchasing a good at a reputable store versus doing so at a flea market. This elucidates [Levmore’s \(1987\)](#) point that since the returns to opportunism vary across environments, the legal rules would also be different.³²

Thinking about opportunism as guile formalized as a strategic manipulation in this paper provides a new way to think about contract design. In legal regimes, contracts are typically required to satisfy the requirement of “good faith and fair dealing”, a concept that traces

³¹See <http://autoweek.com/article/vw-diesel-scandal/germany-seeks-ban-shut-software-wake-vw-diesel-scandal>.

³²In the Volkswagen scandal discussed above, the EPA prosecuted the car manufacturer who then bought back the offending cars from the owner. The law could have been, instead, structured so that the car owners were sued for polluting the environment, and the car owners would in turn sue Volkswagen for the damage they incurred. However, this arrangement clearly would have been more costly for all parties involved.

back to Clause 1 of the 1215 Magna Carta.³³ But the challenge is that the legal framework is often complex and varied. Our model provides a way to formally model “good faith” in contracts while illustrating how a change in the timing of information release can have sharp predictions on the contract form and the scope for opportunism. We hope that this provides a foundation for future research in this area.

REFERENCES

- Abreu, D. (1988). On the theory of infinitely repeated games with discounting. *Econometrica* 56(2), 383–396.
- Baker, G. (1992). Incentive contracts and performance measurement. *Journal of Political Economy* 100, 598–614.
- Baker, G., R. Gibbons, and K. J. Murphy (1994). Subjective performance measures in optimal incentive contracts. *The Quarterly Journal of Economics* 109(439), 1125–1156.
- Banerjee, A. V. and E. Duflo (2000). Reputation effects and the limits of contracting: A study of the Indian software industry. *Quarterly Journal of Economics* 115(3), 989–1017.
- Bull, C. (1987). The existence of self-enforcing implicit contracts. *Quarterly Journal of Economics* 102, 147–159.
- Carmichael, H. L. (1983). The agents-agents problem: Payment by relative output. *Journal of Labor Economics* 1, 50–65.
- Chan, J. and B. Zheng (2011). Rewarding improvements: optimal dynamic contracts with subjective evaluation. *Rand Journal of Economics* 42(4), 758–775.
- Charness, G. and D. Levin (2009). The origin of the winner’s curse: A laboratory study. *American Economic Journal: Microeconomics* 1(1), 207–36.
- Cramer, J. and A. B. Krueger (2016). Disruptive change in the taxi business: The case of uber. Working Paper 22083, National Bureau of Economic Research.
- Crampton, P. (1985). Sequential bargaining mechanisms. In A. E. Roth (Ed.), *Game-Theoretic Models of Bargaining*, Chapter 8, pp. 149–180. Cambridge University Press.
- Darby, M. R. and E. Karni (1973). Free competition and the optimal amount of fraud. *Journal of Law and Economics* 16(April), 67–88.
- Deb, J., J. Li, and A. Mukherjee (2016). Relational contracts with subjective peer evaluations. *The RAND Journal of Economics* 47(1), 3–28.
- Dulleck, U. and R. Kerschbamer (2006). On doctors, mechanics, and computer specialists: the economics of credence goods. *Journal of Economic Literature* 44(1), 5–42.

³³See <https://www.bl.uk/magna-carta/articles/the-clauses-of-magna-carta>.

- Emons, W. (1997). Credence goods and fraudulent experts. *Rand Journal of Economics* 28(1), 107–119. Article.
- Esponda, I. and E. Vespa (2014). Hypothetical thinking and information extraction in the laboratory. *American Economic Journal: Microeconomics* 6(4), 180–202.
- Frank, R. G. and T. G. McGuire (2000). Economics and mental health. In M. V. Pauly, T. G. McGuire, and P. P. Barros (Eds.), *Handbook of Health Economics*, Volume 1, Part B, Chapter 16, pp. 893 – 954. Elsevier.
- Frank, R. H. (1987). If homo economicus could choose his own utility function, would he want one with a conscience? *The American Economic Review* 77(4), 593–604.
- Frank, R. H. (1988). *Passions within Reason*. New York, NY, U.S.A.: W. W. Norton & Company.
- Fuchs, W. (2007). Contracting with repeated moral hazard and private evaluations. *The American Economic Review* 97(4), 1432–1448.
- Gibbons, R. (2005). Incentives between firms (and within). *Management Science* 51(1), 2–17.
- Goffman, E. (1969). *Strategic interaction*. University of Pennsylvania publications in conduct and communication. University of Pennsylvania Press.
- Goodman, S. F. and L. J. Turner (2010). Teacher incentive pay and educational outcomes: Evidence from the new york city bonus program. Discussion Paper, Columbia University.
- Grossman, S. J. and O. D. Hart (1983). An analysis of the principal-agent problem. *Econometrica* 51(1), 7–45.
- Guth, W., R. Schmittberger, and B. Schwarze (1982). An experimental-analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization* 3(4), 367–388.
- Hart, O. D. and J. H. Moore (2007). Contracts as reference points. *Quarterly Journal of Economics* 123(1), 1–48.
- Holmström, B. and P. Milgrom (1987). Aggregation and linearity in the provision of intertemporal incentives. *Econometrica* 55, 303–328.
- Holmström, B. and P. Milgrom (1991). Multi-task principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization* 7, 24–52.
- Kandori, M. (2002). Introduction to repeated games with private monitoring. *Journal of Economic Theory* 102, 1–15.
- Kandori, M. (2011). Weakly belief-free equilibria in repeated games with private monitoring. *Econometrica* 79(3), 877–892.

- Kaur, S. (2014). Nominal wage rigidity in village labor markets. Working Paper 20770, National Bureau of Economic Research.
- Kerr, S. (1975). On the folly of rewarding A, while hoping for B. *Academy of Management Journal* 18(4), 769–783.
- Kranton, R. (1996). Reciprocal exchange: A Self-Sustaining system. *American Economic Review* 86(4), 830–51.
- Lang, M. (2016). Personal correspondence.
- Levin, J. (2003). Relational incentive contracts. *American Economic Review* 93(3), 835–857.
- Levine, D. K. (1998). Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics* 1(3), 593 – 622.
- Levmore, S. (1987). Variety and uniformity in the treatment of the good-faith purchaser. *The Journal of Legal Studies* 16(1), 43–65.
- Li, J. and N. Matouschek (2013). Managing conflicts in relational contracts. *American Economic Review* 103(6), 2328–2351.
- Macchiavello, R. and A. Morjaria (2015). The value of relationships: Evidence from a supply shock to kenyan rose exports. *American Economic Review* 105(9), 2911 – 2945.
- MacLeod, W. B. (2003). Optimal contracting with subjective evaluation. *American Economic Review* 93(1), 216–240.
- MacLeod, W. B. and J. M. Malcomson (1989). Implicit contracts, incentive compatibility, and involuntary unemployment. *Econometrica* 57(2), 447–480.
- MacLeod, W. B. and D. Parent (1999). Job characteristics and the form of compensation. *Research in Labor Economics* 18, 177–242.
- Maestri, L. (2012). Bonus payments versus efficiency wages in the repeated principal-agent model with subjective evaluations. *American Economic Journal: Microeconomics* 4(3), 34–56.
- Malcomson, J. M. (1984). Work incentives, hierarchy, and internal labor markets. *Journal of Political Economy* 92(3), 486–507.
- Malcomson, J. M. (2016). Relational incentive contracts with persistent private information. *Econometrica* 84(1), 317–346.
- Mas, A. (2006). Pay, reference points, and police performance. *Quarterly Journal of Economics* 121(3), 783–821.
- Mas, A. (2008). Labour unrest and the quality of production: Evidence from the construction equipment resale market. *Review of Economic Studies* 75(1), 229–258.
- Maskin, E. and J. Tirole (1999). Unforeseen contingencies and incomplete contracts. *Review of Economic Studies* 66, 83–114.

- Milkovich, G. T., J. M. Newman, and B. Gerhart (2011). *Compensation* (Tenth Edition ed.). Chicago, IL: Irwin.
- Myerson, R. B. (1986). Multistage games with communication. *Econometrica* 54(2), 323–58.
- Myerson, R. B. and M. A. Satterthwaite (1983). Efficient mechanisms for bilateral trading. *Journal of Economic Theory* 29, 265–281.
- Ngangoue, K. and G. Weizsacker (2015). Learning from unrealized versus realized prices. Technical Report 1487, DIW Berlin.
- Prendergast, C. (1999). The provision of incentives in firms. *Journal of Economic Literature* 37(1), 7–63.
- Prendergast, C. (2002). The tenuous trade-off between risk and incentives. *Journal of Political Economy* 110(5), 1071–1102.
- Rahman, D. (2012). But who will monitor the monitor? *The American Economic Review* 102(6), 2767–2797.
- Schmidt, K. M. and M. Schnitzer (1995). The interaction of explicit and implicit contracts. *Economic Letters* 48, 193–199.
- Shapiro, C. and J. E. Stiglitz (1984). Equilibrium unemployment as a worker discipline device. *American Economic Review* 74(3), 433–444.
- Skinner, J. (2012). Causes and consequences of regional variations in health care. In M. V. Pauly, T. G. McGuire, and P. P. Barros (Eds.), *Handbook of Health Economics*, Volume 2, Chapter 2, pp. 45–49. Elsevier B. V.
- Sobel, J. (2006). For better or forever: Formal versus informal enforcement. *Journal of Labor Economics* 24(2), 271–297.
- Telser, L. G. (1980). A theory of self-enforcing agreements. *Journal of Business* 53(1), 27–44. FLA 00219398 University of Chicago Press Copyright 1980 The University of Chicago Press.
- Vigna, S. D. and U. Malmendier (2006). Paying not to go to the gym. *The American Economic Review* 96(3), 694–719.
- Williamson, O. E. (1975). *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: The Free Press.
- Williamson, O. E. (1991). Comparative economic organization: the analysis of discrete structural alternatives. *Administrative Science Quarterly* 36, 269–96.
- Zabojnik, J. (2014). Subjective evaluations with performance feedback. *The RAND Journal of Economics* 45(2), 341–369.

APPENDIX A. OMITTED PROOFS

Proof of Lemma 1.

Proof. Consider a contract where at state $ts \in S^2$, x_{ts} is what the Principal pays out before the conflicts take place, and let $d_{ts}^A > 0$ and $d_{ts}^P > 0$ be the conflicts inflicted by the Agent and the Principal respectively. Under this contract, $c_{ts} = x_{ts} + d_{ts}^A$, $w_{ts} = x_{ts} - d_{ts}^P$ and $\delta_{ts} = c_{ts} - w_{ts} = d_{ts}^A + d_{ts}^P$. Next, consider a contract where $x'_{ts} = x_{ts} + d_{ts}^A$, $d_{ts}^{P'} = d_{ts}^A + d_{ts}^P$, and $d_{ts}^{A'} = 0$; this contract has only the Principal inflicting all the conflicts. It is readily verified that under this contract, $c'_{ts} = x'_{ts} + d_{ts}^{A'} = c_{ts}$, $w'_{ts} = x'_{ts} - d_{ts}^{P'} = w_{ts}$, and $\delta'_{ts} = d_{ts}^{A'} + d_{ts}^{P'} = \delta_{ts}$. Hence, the social loss and both parties' payoffs are left unchanged. A contract that has only the Agent inflicting all the conflicts can be constructed analogously. \square

Proof of Proposition 2.

Proof. The necessity and sufficiency of condition (3.9) has been established in the main text. We are only left prove the properties of $g(\cdot)$. Define the optimal effort function $\Lambda : \mathfrak{R} \rightarrow [0, 1]$ as:

$$(A.1) \quad \Lambda(x) = \begin{cases} (V')^{-1}(x) & , \text{ if } x \in (V'(0), V'(1)), \\ 0 & , \text{ if } x \leq V'(0), \\ 1 & , \text{ if } x \geq V'(1). \end{cases}$$

It is readily observed that $g(x) = \Lambda(x)x - V(\Lambda(x))$. For $x \in (0, V'(1))$, by the envelope theorem, $g'(x) = \Lambda(x) > 0$; and $g''(x) = \Lambda'(x) = \frac{1}{V''(\Lambda(x))} > 0$. \square

Proof of Proposition 3.

Proof. The “only if” direction is trivial. For the “if” direction, take $\vec{w} \in \mathbb{G}(\lambda)$; by Proposition 2, \vec{w} satisfies the guile-free constraint (3.7). Notice that for any $a \in \mathfrak{R}$, $\vec{w} + a\vec{1} \in G(\lambda)$. The Agent's expected payoff from wage term $\vec{w} + a\vec{1}$ is $\vec{\Gamma}(\lambda)\vec{w} + a - V(\lambda)$ and a can be set to satisfy PC (2.8). This implies that the set of guile-free sales contracts $\Psi^S(\lambda)$ is non-empty. As noted in fn. 25, $\Psi^S(\lambda)$ is a convex set. Let \hat{C} be the cost under this contract. Since there are a finite number of states, the set of contracts with costs less than or equal to \hat{C} can be restricted to a bounded set. The linearity of the payoff function, combined with the fact that the feasible set is a closed and convex set, then implies that an optimum (i.e. a feasible contract with the lowest cost) with cost weakly less than \hat{C} exists. \square

Proof of Proposition 4.

Proof. The first point has been established in the main text. For the second point, suppose Π is such that $\vec{\Gamma}^L \Pi \vec{w}^{GF} + \lambda \vec{\tilde{\Gamma}} \Pi \vec{w}^{GF} - V(\lambda) = U^0$ while $\vec{\Gamma}^L \Pi \vec{w}^{GF} \neq \vec{\Gamma}^L \vec{w}^{GF}$. This implies that $\vec{\tilde{\Gamma}} \Pi \vec{w}^{GF} \neq \vec{\tilde{\Gamma}} \vec{w}^{GF}$. Let $\tilde{\lambda}$ be set such that $\tilde{\lambda} = \Lambda \left(\vec{\tilde{\Gamma}} \Pi \vec{w}^{GF} \right)$, where $\Lambda(\cdot)$ is defined in (A.1); $\tilde{\lambda} \neq \lambda$ is the optimal effort level exerted by the Agent under reporting strategy Π . Hence we have

$$\begin{aligned} 0 &= \left[\vec{\Gamma}^L \Pi \vec{w}^{GF} + \lambda \vec{\tilde{\Gamma}} \Pi \vec{w}^{GF} - V(\lambda) \right] - \left[\vec{\Gamma}^L \vec{w}^{GF} + \lambda \vec{\tilde{\Gamma}} \vec{w}^{GF} - V(\lambda) \right], \\ &< \left[\vec{\Gamma}^L \Pi \vec{w}^{GF} + \tilde{\lambda} \vec{\tilde{\Gamma}} \Pi \vec{w}^{GF} - V(\tilde{\lambda}) \right] - \left[\vec{\Gamma}^L \vec{w}^{GF} + \lambda \vec{\tilde{\Gamma}} \vec{w}^{GF} - V(\lambda) \right]. \end{aligned}$$

□

Proof of Proposition 5.

Proof. It is readily observed that in the expert-Agent environment, $\Gamma_{ts}(\lambda) = \left[\lambda \beta_s^H + (1 - \lambda) \beta_s^L \right] r_{ts}$. Hence, the conditional probability of the Principal receiving signal t given that the Agent has put in effort λ and observed signal s is $Pr[t|s, \lambda] = \frac{\Gamma_{ts}(\lambda)}{\sum_{t' \in S} \Gamma_{t's}(\lambda)} = r_{ts}$ which is independent of λ . This implies that the Agent's truthful-reporting constraint ATR (3.4) here is:

$$(A.2) \quad \left(p_s - \sum_{t \in S} r_{ts} \delta_{ts}^S \right) - \left(p_{s'} - \sum_{t \in S} r_{ts} \delta_{ts}^{S'} \right) \geq 0, \quad \forall s, s' \neq s,$$

which is also independent of the Agent's effort. Let \vec{w} by the wage vector for a good faith sales contract ψ^{GF} and suppose, for a contradiction, that there exists a reporting strategy $\Pi \in Z$ that violates the guile-free constraint. Let $\lambda^\Pi = \Lambda \left(\vec{\tilde{\Gamma}} \Pi \vec{w} \right)$, as defined in (A.1), which is the optimal effort level that attains the maximum expected payoff for the Agent under reporting strategy Π . Let $u^A(\lambda', \Pi | \psi)$ be the Agent's ex-ante expected payoff for choosing effort λ' together with reporting strategy Π . Since adhering to the contract terms gives the Agent an expected payoff of U^0 , we have:

$$(A.3) \quad U^0 < u^A \left(\lambda^\Pi, \Pi | \psi^{GF} \right)$$

$$(A.4) \quad \leq u^A \left(\lambda^\Pi, I | \psi^{GF} \right)$$

$$(A.5) \quad \leq u^A \left(\lambda, I | \psi^{GF} \right) \\ = U^0,$$

where the first inequality in (A.3) follows from the assumption that the guile-free constraint is violated for Π , the second inequality in (A.4) follows from (A.2) that the Agent is always weakly better off by truthfully reporting at every signal regardless of the exerted effort, and the third inequality in (A.5) follows from ICE (3.3) under truthful-reporting. We thus have a contradiction. \square

Proof of Proposition 7.

Proof. Denote $\hat{\gamma}_t = \sum_{s \in S} \hat{\Gamma}_{ts}$, and let $\vec{\hat{\gamma}}$ be the $1 \times n$ vector of $\hat{\gamma}_t$. Suppose an optimal authority contract exists. The contract satisfies ICE (3.3) which implies that $V'(\lambda) = \sum_{ts \in S^2} w_{ts} = \sum_{t \in S} \hat{\gamma}_t p_t^A$. Since $V'(\lambda) > 0$, there must exist some t such that $\hat{\gamma}_t \neq 0$.

Conversely, suppose that $\hat{\gamma}_t \neq 0$ for some t . Since $\sum_{t \in S} \hat{\gamma}_t = 0$, there exists t' such that $\hat{\gamma}_{t'} > 0$. Consider a price vector \vec{p}^0 such that $p_t^0 = 0 \forall t \neq t'$, and $p_{t'}^0 = b = \frac{V'(\lambda)}{\hat{\gamma}_{t'}}$. Hence $\sum_{t \in S} \hat{\gamma}_t p_t^0 = V'(\lambda)$. Let \vec{p}^1 be a price vector in which $p_t^1 = p_t^0 + a$. Notice $\sum_{t \in S} \hat{\gamma}_t p_t^1 = \sum_{t \in S} \hat{\gamma}_t p_t^0 = V'(\lambda)$. Hence \vec{p}^1 satisfies ICE (3.3), and a can be set such that \vec{p}^1 also satisfies PC (2.8). Next, $\forall s \in S$, set $\delta_{ts}^A = b \forall t \neq t'$, and $\delta_{t's}^A = 0$. This implies that $c_{ts} = p_t^1 + \delta_{ts}^A = b \forall t, s \in S$. Hence PTR (5.3) is trivially satisfied. This implies that the feasible set for the contracting problem for the optimal authority contract is non-empty. The Principal's total expected cost under this contract is b . By the same argument as in Proposition 3, an optimum with cost weakly less than b exists. \square

DEPARTMENT OF ECONOMICS, COLUMBIA UNIVERSITY. E-MAIL: BENTLEY.MACLEOD@COLUMBIA.EDU.
E-MAIL: TECK.TAN@COLUMBIA.EDU