Vierø, Marie-Louise

Working Paper
Bait contracts

Queen's Economics Department Working Paper, No. 1212

Provided in Cooperation with:
Queen's University, Department of Economics (QED)

Suggested Citation: Vierø, Marie-Louise (2009) : Bait contracts, Queen's Economics Department Working Paper, No. 1212, Queen's University, Department of Economics, Kingston (Ontario)

This Version is available at:
http://hdl.handle.net/10419/67769

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Bait Contracts

Marie-Louise Vierø
Queen’s University

Department of Economics
Queen’s University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

8-2009
Bait Contracts*

Marie-Louise Vierø†
Queen’s University
January 17, 2010

Abstract

The granting of stock options to employees who have negligible impact on company performance intuitively violates Holmström’s (1979) sufficient statistic result. This paper revisits the sufficient statistic question of when to condition a contract on an outside signal in a principal-agent model in which I introduce imprecise (or vague) information. The paper applies a choice theoretic framework introduced in Olszewski (2007) and Ahn (2008) and extended by Vierø (2009a), who denoted it vague environments. I show that if the signal is vague, Holmström’s result can be overturned.

Keywords: contracts, vagueness, optimism, incentives, signals, stock options
JEL classification: D82, D80, D20, D86

1 Introduction

In the 1990s, dot-com companies began offering stock options to a wide range of their employees, including rank-and-file employees, and the use of stock options as part of the compensa-

*I thank Daron Acemoglu, Levon Barseghyan, Larry Blume, Amrita Dhillon, David Easley, Edi Karni, George Mailath, Morten Nielsen, Jan Zabojnik, and seminar participants at the 2007 SAET conference, the 2008 FUR conference, University of Heidelberg, The Institute of Advanced Studies in Vienna, Johns Hopkins University, the Norwegian School of Economics, Purdue University, Santa Clara University, and University of Warwick for comments and suggestions.

†Please address correspondence to: Marie-Louise Vierø, Department of Economics, Dunning Hall Room 316, Queen’s University, 94 University Avenue, Kingston, Ontario K7L 3N6, Canada; phone: (+1) 613-533-2292; e-mail: viero@econ.queensu.ca
tion package for non-executive employees quickly spread to other publicly held companies.\footnote{For example, in a sample of 889 firms, Bergman and Jenter (2007) find that employees ranking below the top-five executives receive 71\% of the options granted.} While executives have a direct influence on the company’s performance, this is not so for e.g. the individual janitor, who has negligible impact on the company’s stock price. Hence, the usual explanation that stock options align the incentives of the employee with those of the shareholders does not appear to apply. The present paper provides an explanation for granting stock options to employees who have negligible impact on company performance, a practice which intuitively violates Holmström’s (1979) sufficient statistic result.

Something that characterized the dot-com industry in the 1990s was that it was a relatively new industry facing a lot of uncertainty; uncertainty about which little was known. This suggests that imprecise (or vague) information could be the key to understanding why companies offer stock options to a wider set of employees. In a contracting framework, this paper shows that vague information does indeed play a key role in explaining this phenomenon. The intuition behind the result is the same as the argument usually given in the context of the dot-com companies: that they offered stock options to compensate for lower salaries.\footnote{While later arguments have also included giving workers a sense of ownership, the original argument of lower salaries is still important.}

This paper applies a choice theoretic framework, which is exactly one of vague or imprecise information, to a canonical principal-agent model with hidden information. The environment is vague because the decision makers are assumed to only know a set of possible probability distributions, or lotteries, over outcomes, rather than a precise probability distribution as is usually assumed. I use Vierø’s (2009a) Optimism-Weighted Subjective Expected Utility (OWSEU) representation of preferences in vague environments, which in the one-state version applied here also corresponds to one of the representations in Olszewski (2007).\footnote{Ahn (2008) considers a different axiomatization and representation for the one-state version.} In the present context this representation characterizes a decision maker by his Bernoulli-utility and his optimism, and models him as if he evaluates an act (a contract in the present context) by computing the usual von Neumann-Morgenstern utility of the best lottery and of the worst lottery in the set of lotteries and weighting them together, where the weight on the best lottery can be interpreted as the decision maker’s level of optimism.\footnote{In the following I use the terminology from Vierø (2009a) and refer to the theory therein.}
The distinctive feature of the model in the present paper is the presence of a vague public signal. The main result of the paper is to show that even if the public signal is uninformative about the private information of the agent, the principal can benefit from conditioning the contract on the vague public signal. Hence, orthogonal signals can be of value if they are vague. I denote contracts that condition on the signal 'bait contracts'.

In a standard model with no vagueness, Holmström (1979) considers the conditions under which a principal can improve upon a contract by conditioning on an outside signal. He shows that it is optimal to make the contract contingent on the outside signal if and only if the signal is not orthogonal to the directly payoff relevant variables of interest. The result in the present paper, that it can be optimal to condition on an outside vague signal even if this signal is orthogonal to the unobserved variables that directly affect payoffs, is therefore in contrast with Holmström’s result. When bait contracts are optimal, the principal can exploit the presence of optimism by deliberately introducing vagueness into the contract. The optimality of bait contracts provides an explanation for the granting of stock options to employees who have negligible impact on the stock price.

By conditioning on the signal, the principal can influence the agent’s weight, or emphasis, on the different final scenarios. Thus, conditioning on the signal endogenously cause the principal and agent to have heterogeneous beliefs. As a consequence, the contract fulfills two purposes. On one hand, it serves the usual purpose of ensuring participation of and providing incentives for the agent. On the other hand, the parties are betting on their differences in emphasis through the contract. When the parties are sufficiently optimistic, the principal can exploit that they emphasize different final scenarios, and the principal can optimally generate such a difference in emphasis by offering a bait contract.

An important thing to note is that no heterogeneity in optimism between the contracting parties is needed for bait contracts to be optimal. Note also that the notion of optimism in the present paper is different from the notion of overconfidence in people’s own abilities that we see in e.g. Benabou and Tirole (2002). In the present paper, optimism is a consistent feature of the individual’s personality, which affects his or her perspective on life in general.

Vierø (2009b) considers contracting between risk neutral parties when the contracting

---

5The intuition for the use of the term ‘bait contract’ should become clear below.

6Andersen, Fountain, Harrison, and Rutström (2009) provide experimental evidence of the relevance of optimism.
environment itself is vague, and shows that the presence of vagueness often leads to the
standard ‘sell the firm to the agent’ contract being suboptimal. In the present paper, on the
other hand, the contracting environment itself is precise, while the outside signal is vague.
Mukerji (1998) and Rigotti (2006) analyze other contracting problems in non-standard choice
theoretic settings. Mukerji (1998) considers a moral hazard problem and shows that ambigu-
ity aversion among the parties can rationalize incomplete contracts. Rigotti (2006) considers
a principal-agent model with moral hazard in which the agent has incomplete preferences.

The present paper is also related to a group of papers that consider contracting when
the parties have heterogeneous beliefs. These include Adrian and Westerfield (2009) and
Carlier and Renou (2005, 2006). When beliefs are heterogeneous, the parties also want to
place side-bets on the resolution of uncertainty, but there is no possibility for the principal
to influence the agent’s weight on the different final scenarios. With precise information and
heterogeneous beliefs, all differences between the contracting parties are exogenous.

The paper is organized as follows: Section 2 presents the model with vagueness. Section
3 contains the bait contracts result. Section 4 concludes.

2 Model

Consider the canonical principal-agent problem with hidden information. A risk neutral
principal (she) wants to hire a risk averse agent (he) to complete a task. It is assumed that
the agent’s utility depends on a variable, measuring how well suited to the required task he
will find himself, the value of which is realized after the contract is signed. For convenience,
I will refer to this variable as the agent’s efficiency level, but it could be interpreted in a
variety of ways. Suppose the agent’s effort can be measured by a one-dimensional variable
e ∈ [0, ∞). The principal’s gross profit is a function of the agent’s effort, π(e), with π(0) = 0,
π'(e) > 0, and π''(e) < 0 ∀e. Her Bernoulli utility function is given by her net profits,

\[ u_P(w, e) = \pi(e) - w, \]

where w denotes the wage she pays to the agent.

The agent’s Bernoulli utility function depends on his wage w, how much effort he chooses
to exert, and his efficiency x, which affects how much disutility, denoted g(e, x), he expe-
riences from effort. It is assumed that there are only two possible values of x: the agent
will be either of high-efficiency type $x_H$ or of low-efficiency type $x_L$. The efficiency level is unobservable to the principal, while effort is assumed to be observable and contractible.

Assume further that the agent’s Bernoulli utility function is of the form

$$u_A(w,e,x) = v(w - g(e,x)),$$

with $v'(\cdot) > 0$ and $v''(\cdot) < 0$.

The disutility $g(e,x)$ is assumed to satisfy the following standard conditions: $g(0,x_H) = g(0,x_L) = g_e(0,x_H) = g_e(0,x_L) = 0$, such that the agent suffers no disutility if he does not exert any effort, $g_e(e,x) > 0 \forall e > 0$ and $g_{ee}(e,x) > 0 \forall e$, such that his disutility from effort is increasing at an increasing rate, and $g(e,x_L) > g(e,x_H) \forall e > 0$ and $g_e(e,x_L) > g_e(e,x_H) \forall e > 0$, such that his disutility and marginal disutility from effort are higher if he is of low-efficiency type. Finally, let $\bar{u}$ denote the agent’s reservation utility.

Let $p = (p_H, p_L)$ denote the probability distribution over the agent’s types. That is, the agent will be of type $x_i$ with probability $p_i$. This probability is known to both parties to the contract. Hence, the contracting environment itself is precise, i.e. there is no vagueness about the agent’s type.

Suppose there is a publicly observable outside signal, which can take values $y_H$ or $y_L$. The contracting parties do not know the precise probability with which the signal will be of value $y_H$. Instead they only know a possible set $Q = \{q = (q_H, q_L) : q_H \in [c,d], q_L = 1 - q_H\}$ of this probability distribution. That is, the outside signal is vague, with the probability of $y_H$ being $q_H \in [c,d] \subseteq [0,1]$. The parties therefore have common but vague knowledge of the probability of the signal being $y_H$.

I further assume that the signal is orthogonal to the agent’s type, i.e. the probability of realizing the pair $(x_i, y_j)$, $i, j = H, L$, is the product of the relevant marginal probabilities.

A contract consists of a menu of wage-effort pairs. The principal can choose whether or not to write a contract that is contingent on the vague signal. I denote contracts that do condition on the signal ‘bait contracts’. Such a bait contract will consist of wage-effort pairs for each type for each value of the signal:

$$(e(x_H,y_H), w(x_H,y_H), e(x_H,y_L), w(x_H,y_L), e(x_L,y_H), w(x_L,y_H), e(x_L,y_L), w(x_L,y_L))$$

$\equiv (e_{HH}, w_{HH}, e_{HL}, w_{HL}, e_{LH}, w_{LH}, e_{LL}, w_{LL})$.

Given the probability distribution $p$ over types and the set of probability distributions $Q$ over signals, each contract $(e_{HH}, w_{HH}, e_{HL}, w_{HL}, e_{LH}, w_{LH}, e_{LL}, w_{LL})$ induces an act $h =$
which is a set of lotteries over final outcomes 

\[ z = (w, e, x) \]

The information structure is an example of a vague environment as defined in Vierø (2009a), who directly allows the decision environment to be imprecise. The present paper applies a one-state version of her model, which also corresponds to the model in Olszewski (2007), in which acts are sets of lotteries over outcomes.\(^7\) Vierø (2009a) shows that if preferences satisfy the standard axioms for subjective expected utility, properly expanded to her generalized domain, plus two additional axioms that are natural extensions of the standard axioms when considering vague environments, then decision makers can be modeled as Optimism-Weighted Subjective Expected Utility (OWSEU) maximizers. I assume that the contracting parties have such OWSEU-preferences.

In the present context the assumption of OWSEU-preferences implies that both the principal and the agent maximize utility of the following form:

\[
\text{OWSEU}_k(h) = \alpha_k \left[ \sum_{j \in \{H,L\}} \sum_{i \in \{H,L\}} q_{k,j}(h)p_i u_k(z_{ij}) \right] + (1 - \alpha_k) \left[ \sum_{j \in \{H,L\}} \sum_{i \in \{H,L\}} q_{k,j}(h)p_i u_k(z_{ij}) \right],
\]

where each sum is over the support of the relevant lottery, \( j \) indexes the value of the signal, \( i \) indexes the agent’s type, \( k \in \{P,A\} \), \( u_k \) is \( k \)'s Bernoulli utility function defined over outcomes \( z_{ij} = (w_{ij}, e_{ij}, x_i) \), and \( \alpha_k \in (0, 1) \) is a parameter that captures \( k \)'s degree of optimism. Finally, \( q_{k,H}(h) = (\overline{q}_{k,H}(h), \overline{q}_{k,L}(h)) \) and \( q_{k,L}(h) = (\underline{q}_{k,H}(h), \underline{q}_{k,L}(h)) \) are, respectively, the best and worst probability distributions in the set \( Q \) from \( k \)'s point of view. That is,

\[
\overline{q}_{k}(h) = \arg \max_{q \in Q} \sum_{j \in \{L,H\}} \sum_{i \in \{L,H\}} q_j p_i u_k(z_{ij})
\]

and

\[
\underline{q}_{k}(h) = \arg \min_{q \in Q} \sum_{j \in \{L,H\}} \sum_{i \in \{L,H\}} q_j p_i u_k(z_{ij}).
\]

Since \( x \) and \( y \) are orthogonal to each other, there is a one-to-one correspondence between these best and worst marginal probabilities and the best and worst overall probabilities over final outcomes. One interpretation of equation (1) is that to evaluate an act \( h \), the decision

\(^7\)The model in Vierø (2009a) is actually a model of any finite number of states with acts that are mappings from states into sets of lotteries over outcomes.
maker computes the von Neumann-Morgenstern utility of the best lottery and of the worst lottery in the set of lotteries. These are given in the first and second sets of square brackets in (1), respectively. The decision maker then weighs them together with weight $\alpha_k$ on the best lottery. The weight $\alpha_k$ can thus be interpreted as the decision maker’s level of optimism.

Contracting is assumed to take place ex-ante, that is, before the agent knows his type. Ex-ante contracting has two stages: the agent first agrees to a menu of wage-effort pairs, and then, once he learns his type, selects one of the wage-effort pairs in the menu. I assume that the principal is unable to observe the agent’s efficiency level at any point in time. Hence, an informational asymmetry arises ex-post, after the contract is signed, while at the time of contracting there is no asymmetric information.

It is important to note that which lotteries are best and worst depend on the contract offered. Therefore, the contract offered endogenously determines the beliefs of the agent and principal. Consequently, the agent and the principal may endogenously have heterogeneous beliefs. This is the key consequence of vagueness and a driving force behind the optimality of bait contracts.

3 Bait contracts

I now show that for some principal-agent matches, the principal will write deliberately vague contracts. That is, the principal will optimally introduce vagueness into a contracting situation where there is otherwise no vagueness, since the principal will optimally choose to condition the contract on a vague signal that does not contain information about the agent’s type. This implies that vague signals can be valuable, even if they are orthogonal to the unobserved variables that directly affect the parties’ payoffs.

In a standard model with no vagueness, Holmström (1979) considers the question of when an outside signal can be used to improve upon a contract. He shows that it is optimal to make the contract contingent on the outside signal if and only if the signal is not orthogonal to the directly payoff relevant variables of interest. Holmström obtains his result in a moral hazard model, but a similar result can be derived in a model with adverse selection and precise information.\footnote{In fact it follows by setting $c = d$ below that the principal will not condition on an orthogonal signal if} I now show that Holmström’s result can be overturned in some cases.
if the outside event adds vagueness to the contracting situation.

The principal can choose whether or not to write a contract that is contingent on the vague signal, i.e. whether or not to write a bait contract. Since the signal is publicly observable, conditioning on it will not lead to any additional informational asymmetries, but a bait contract does have to be incentive compatible given each value of the signal. A contract that does not condition on the signal has \( e(x_H, y_H) = e(x_H, y_L) \equiv e_H \), \( e(x_L, y_H) = e(x_L, y_L) \equiv e_L \), \( w(x_H, y_H) = w(x_H, y_L) \equiv w_H \), and \( w(x_L, y_H) = w(x_L, y_L) \equiv w_L \). Hence, it is a special case of conditioning.

The following lemma shows that the revelation principle holds in vague environments and is proven in Vierø (2009b).

**Lemma 1.** In a vague environment, any general incentive compatible contract can be implemented with a truthful revelation mechanism.

Given Lemma 1, the principal’s problem of deciding whether to write a vague contract is given by

\[
\max_{\substack{w_{HH}, e_{HH} \geq 0 \\ w_{HL}, e_{HL} \geq 0 \\ w_{LL}, e_{LL} \geq 0}} \left( \alpha_P q_{P,H} + (1 - \alpha_P) q_{P,H}L \right) p_H \left( \pi(e_{HH}) - w_{HH} \right) + p_L \left( \pi(e_{LL}) - w_{LL} \right) \\
+ \left( \alpha_P q_{P,L} + (1 - \alpha_P) q_{P,L} \right) p_H \left( \pi(e_{HL}) - w_{HL} \right) + p_L \left( \pi(e_{LL}) - w_{LL} \right)
\]

subject to

\[
\bar{\pi} \leq \left( \alpha_A \bar{q}_{A,H} + (1 - \alpha_A) \bar{q}_{A,L} \right) p_H v(w_{HH} - g(e_{HH}, x_H)) + p_L v(w_{HL} - g(e_{HL}, x_L)) \\
+ \left( \alpha_A \bar{q}_{A,L} + (1 - \alpha_A) \bar{q}_{A,L} \right) p_H v(w_{HL} - g(e_{HL}, x_H)) + p_L v(w_{LL} - g(e_{LL}, x_L)),
\]

\[
v(w_{HH} - g(e_{HH}, x_H)) \geq v(w_{HH} - g(e_{HH}, x_H)), \quad (IC_{HH})
\]
\[
v(w_{HL} - g(e_{HL}, x_L)) \geq v(w_{HL} - g(e_{HL}, x_L)), \quad (IC_{HL})
\]
\[
v(w_{LL} - g(e_{LL}, x_L)) \geq v(w_{LL} - g(e_{LL}, x_L)), \quad (IC_{LL})
\]

where \( \bar{q}_k(h) = (\bar{q}_{k,H}(h), \bar{q}_{k,L}(h)) \) and \( q_k(h) = (q_{k,H}(h), q_{k,L}(h)) \) are given in (2) and (3), respectively, and denote the best and worst probability distributions in \( Q \) from \( k \)’s point of the environment is precise.
view given the contract. For the rest of the paper, I suppress the dependency on \( h \) to ease notation and simply write \( \eta_{k,j} \) and \( \eta_{k,j} \), but the dependency is nonetheless there.

When conditioning on the vague signal, the principal creates a disagreement among the parties about which realization of the signal is the best. Thus, the best lottery in \( Q \) from the principal’s point of view is the worst lottery from the agent’s point of view and vice versa. This generally creates an endogenous difference in emphasis on the possible final scenarios between the principal and the agent that would otherwise not be there. Consequently, the contract fulfills two purposes. Not only does it serve the usual purpose of ensuring participation of and providing incentives for the agent, but the parties are also placing side-bets on their differences in emphasis through the contract. When the parties are sufficiently optimistic, the principal can exploit that they emphasize different final scenarios, and therefore it is worthwhile for her to generate such a difference. Thus, with sufficient optimism, the principal can use the vague signal to her advantage.

The following theorem shows that bait contracts can be optimal, i.e. that there exist principals and agents for which the principal will condition on the vague signal.

**Theorem 1 (Existence of bait contracts).** Let \( p = (p_H, p_L) \) denote the probability distribution over the agent’s type \( x \), i.e. there is no vagueness about the unobserved variable that directly affects the parties’ payoffs. Let the set of possible probability distributions over the outside signal \( y \) be given by \( Q \). Furthermore, let \( x \) and \( y \) be orthogonal to each other. Then the following two statements are equivalent:

i. There exists a set of principals and agents characterized by \( \alpha_P, \pi(\cdot), \alpha_A, v(\cdot), g(e,x) \), respectively, for which the optimal contract has \((e_{HH}, w_{HH}, e_{LH}, w_{LH}) \neq (e_{HL}, w_{HL}, e_{LL}, w_{LL})\).

ii. The outside signal is vague, i.e. \( c < d \) in \( Q \).

That is, if and only if the outside signal is vague, there exists a set of principals and agents for which the principal will optimally choose to make the contract contingent on the realization of the outside signal, even when this signal is uninformative about the unobserved variable that directly affects the parties’ payoffs.

**Proof:** I first show that i. implies ii. by showing the contrapositive. This follows simply by noting that if ii. is not true (i.e., if \( c = d \)) we are in the standard model with precise information, which implies that i. is not true.
I now show that \( ii. \) implies \( i. \) Suppose for simplicity that \( \pi = 0. \) The Lagrangian for the principal’s problem is

\[
\mathcal{L} = (\alpha P \bar{q}_P + (1 - \alpha P) q_P) \left( p(\pi(e_{HH}) - w_{HH}) + (1 - p)(\pi(e_{LL}) - w_{LL}) \right) \\
+ (\alpha P(1 - \bar{q}_P) + (1 - \alpha P)(1 - q_P)) \left( p(\pi(e_{HL}) - w_{HL}) + (1 - p)(\pi(e_{LL}) - w_{LL}) \right) \\
+ \gamma \left[ (\alpha A \bar{q}_A + (1 - \alpha A) q_A) \left( p v(w_{HH} - g(e_{HH}, x_H)) + (1 - p)v(w_{HL} - g(e_{HL}, x_L)) \right) \\
+ (\alpha A(1 - \bar{q}_A) + (1 - \alpha A)(1 - q_A)) \left( p v(w_{HL} - g(e_{HL}, x_H)) + (1 - p)v(w_{LL} - g(e_{LL}, x_L)) \right) \right] \\
+ \lambda_{HH} [v(w_{HH} - g(e_{HH}, x_H)) - v(w_{HL} - g(e_{HL}, x_H))] \\
+ \lambda_{HL} [v(w_{HL} - g(e_{HL}, x_H)) - v(w_{LL} - g(e_{LL}, x_L))] \\
+ \lambda_{LL} [v(w_{LL} - g(e_{LL}, x_L)) - v(w_{HL} - g(e_{HL}, x_L))].
\]

The first-order conditions for the problem are

\[
p(\alpha P \bar{q}_P + (1 - \alpha P) q_P) p'(e_{HH}) - \gamma p(\alpha A \bar{q}_A + (1 - \alpha A) q_A) v'(w_{HH} - g(e_{HH}, x_H)) g_e(e_{HH}, x_H) \\
- \lambda_{HH} v'(w_{HH} - g(e_{HH}, x_H)) g_e(e_{HH}, x_H) + \lambda_{HL} v'(w_{HL} - g(e_{HL}, x_L)) g_e(e_{HL}, x_L) \leq 0,
\]

\[
-\lambda_{HH} v'(w_{HH} - g(e_{HH}, x_H)) - \lambda_{HL} v'(w_{HL} - g(e_{HL}, x_L)) = 0,
\]

\[
(1 - p)(\alpha P \bar{q}_P + (1 - \alpha P) q_P) p'(e_{LL}) - \gamma (1 - p)(\alpha A \bar{q}_A + (1 - \alpha A) q_A) v'(w_{HL} - g(e_{HL}, x_L)) g_e(e_{HL}, x_L) \\
+ \lambda_{HH} v'(w_{HH} - g(e_{HH}, x_H)) g_e(e_{HH}, x_H) - \lambda_{HL} v'(w_{HL} - g(e_{HL}, x_L)) g_e(e_{HL}, x_L) \leq 0,
\]

\[
-(1 - p)(\alpha P \bar{q}_P + (1 - \alpha P) q_P) + \gamma (1 - p)(\alpha A \bar{q}_A + (1 - \alpha A) q_A) v'(w_{HL} - g(e_{HL}, x_L)) g_e(e_{HL}, x_L) \\
- \lambda_{HH} v'(w_{HL} - g(e_{HL}, x_L)) - \lambda_{HL} v'(w_{LL} - g(e_{LL}, x_L)) = 0,
\]

\[
p(\alpha P(1 - \bar{q}_P) + (1 - \alpha P)(1 - q_P)) p'(e_{HL}) \\
- \gamma p(\alpha A(1 - \bar{q}_A) + (1 - \alpha A)(1 - q_A)) v'(w_{HL} - g(e_{HL}, x_H)) g_e(e_{HL}, x_H) \\
- \lambda_{HL} v'(w_{HL} - g(e_{HL}, x_H)) g_e(e_{HL}, x_H) + \lambda_{LL} v'(w_{LL} - g(e_{LL}, x_L)) g_e(e_{HL}, x_L) \leq 0,
\]

\[
-\lambda_{HH} v'(w_{HH} - g(e_{HH}, x_H)) + \lambda_{HL} v'(w_{HL} - g(e_{HL}, x_H)) + \lambda_{LL} v'(w_{LL} - g(e_{LL}, x_L)) = 0,
\]

\[
0.
\]
\[
(1 - p)(\alpha_p(1 - \overline{q}_p) + (1 - \alpha_p)(1 - \underline{q}_p))\pi'(e_{LL})
- \gamma(1 - p)(\alpha_A(1 - \overline{q}_A) + (1 - \alpha_A)(1 - \underline{q}_A))v'(w_{LL} - g(e_{LL}, x_L))g_e(e_{LL}, x_L)
+ \lambda_{HL}v'(w_{LL} - g(e_{LL}, x_H))g_e(e_{LL}, x_H) - \lambda_{LL}v'(w_{LL} - g(e_{LL}, x_L))g_e(e_{LL}, x_L) \leq 0,
\]
\[
- (1 - p)(\alpha_P(1 - \overline{q}_P) + (1 - \alpha_P)(1 - \underline{q}_P))
+ \gamma(1 - p)(\alpha_A(1 - \overline{q}_A) + (1 - \alpha_A)(1 - \underline{q}_A))v'(w_{LL} - g(e_{LL}, x_L))
- \lambda_{HL}v'(w_{LL} - g(e_{LL}, x_H)) - \lambda_{LL}v'(w_{LL} - g(e_{LL}, x_L)) = 0,
\]
\[
\overline{u} \leq (\alpha_A\overline{q}_A + (1 - \alpha_A)\underline{q}_A)\left( p \left( v(w_{HH} - g(e_{HH}, x_H)) + (1 - p)v(w_{LU} - g(e_{LU}, x_L)) \right) + (\alpha_A(1 - \overline{q}_A) + (1 - \alpha_A)(1 - \underline{q}_A))p \left( v(w_{HL} - g(e_{HL}, x_H)) + (1 - p)v(w_{LL} - g(e_{LL}, x_L)) \right) \right),
\]
\[
(1C_{HH})
\]
\[
(1C_{LL})
\]
\[
(1C_{HL})
\]
\[
(1C_{LH})
\]
where (4), (6), (8), (10), (PC), (1C_{HH}), (1C_{LH}), (1C_{HL}), and (1C_{LL}) hold with equality if, respectively, \(e_{HH}, e_{LU}, e_{HL}, e_{LL}, \gamma, \lambda_{HH}, \lambda_{LU}, \lambda_{HL}, \) and \(\lambda_{LL}\) are strictly greater than zero.

It follows from (5) and (7) that \(\gamma^* > 0\). Furthermore, since \(\pi'(0) > 0\) and \(g_e(0, x_L) = g_e(0, x_H) = 0\), it follows from (4), (6), (8), and (10), respectively, that \(e'_{HH} > 0\), \(e'_{LU} > 0\), \(e'_{HL} > 0\), and \(e'_{LL} > 0\). Thus, we have equality in (PC), (4), (6), (8), and (10).

The following constellations of the Lagrange multipliers,
\[
\lambda_{HL} = 0 \text{ and } \lambda_{LL} = 0,
\]
\[
\lambda_{HH} = 0 \text{ and } \lambda_{LU} = 0,
\]
\[
\lambda_{HH} > 0 \text{ and } \lambda_{LU} > 0,
\]
\[
\lambda_{HL} > 0 \text{ and } \lambda_{LL} > 0,
\]
\[
\lambda_{HL} = 0 \text{ and } \lambda_{LL} > 0,
\]
and
\[
\lambda_{HH} = 0 \text{ and } \lambda_{LU} > 0,
\]
all lead to contradictions. When \( \lambda_{HL} = \lambda_{LL} = 0 \), (9) and (11) imply that \( w_{LL} - g(e_{LL}, x_L) = w_{HL} - g(e_{HL}, x_H) \), which violates \((IC_{HL})\) and \((IC_{LL})\). A similar argument gives a contradiction when \( \lambda_{HH} = \lambda_{LL} = 0 \).

When \( \lambda_{HH} > 0 \) and \( \lambda_{LH} > 0 \), \((IC_{HH})\) and \((IC_{LH})\) give that \( e_{HH} = e_{LH} \) and \( w_{HH} = w_{LH} \). By (4), (5), (6), and (7), it then follows that \( \pi'(e_{HH}) - g_e(e_{HH}, x_H) < \pi'(e_{HH}) - g_e(e_{HH}, x_L) \), which contradicts the assumptions on the function \( g(e, x) \). A similar argument can be used to show that \( \lambda_{HL} > 0 \) and \( \lambda_{LL} > 0 \) leads to a contradiction.

Finally, when \( \lambda_{HL} = 0 \) and \( \lambda_{LL} > 0 \), (9), (11), and \((IC_{LL})\) can be combined to solve for \( \lambda_{LL} \) as a function of \( e_{HL} \) and \( w_{HL} \). The condition that \( \lambda_{LL} > 0 \) then gives that \( w_{HL} - g(e_{HL}, x_H) < w_{LL} - g(e_{LL}, x_L) \), which violates \((IC_{HL})\) and \((IC_{LL})\). By a similar argument, \( \lambda_{HH} = 0 \) and \( \lambda_{LH} > 0 \) leads to a contradiction.

This leaves one case to investigate, namely

\( \lambda_{HH} > 0, \lambda_{LH} = 0, \lambda_{HL} > 0, \) and \( \lambda_{LL} = 0 \).

In this case, (4) and (5) imply that

\[
\pi'(e_{HH}^*) = g_e(e_{HH}^*, x_H),
\]

while (8) and (9) imply that

\[
\pi'(e_{HL}^*) = g_e(e_{HL}^*, x_H).
\]

Thus, \( e_{HH}^* = e_{HL}^* \). Also, by (6) and (7),

\[
(1 - p)(\alpha P \bar{q}_P + (1 - \alpha P)q_P) \left( \pi'(e_{HL} - g(e_{HL}, x_L)) \right) = \lambda_{HH} v'(w_{HL} - g(e_{HL}, x_H)(g_e(e_{HL}, x_L) - g_e(e_{HL}, x_H))),
\]

and by (10) and (11),

\[
(1 - p)(\alpha P(1 - \bar{q}_P) + (1 - \alpha P)(1 - q_P)) \left( \pi'(e_{HL} - g(e_{HL}, x_L)) \right) = \lambda_{HL} v'(w_{LL} - g(e_{LL}, x_H)(g_e(e_{LL}, x_L) - g_e(e_{LL}, x_H))).
\]

It follows from \((IC_{HH})\) and \((IC_{HL})\) that

\[
w_{HH} - g(e_{HH}, x_H) = w_{LH} - g(e_{LH}, x_H)
\]

and

\[
w_{HL} - g(e_{HL}, x_H) = w_{LL} - g(e_{LL}, x_H).
\]
Similarly, (9) and \((IC_{HH})\) imply that
\[
\lambda_{HH} = \frac{p(\alpha p \overline{q}_P + (1 - \alpha p)q_P)}{v'(w_{LH} - g(e_{LH}, x_H))} - \gamma p(\alpha A \overline{q}_A + (1 - \alpha A)q_A),
\]
which together with (7) implies that
\[
\gamma = \frac{\alpha p \overline{q}_P + (1 - \alpha p)q_P}{p(\alpha A \overline{q}_A + (1 - \alpha A)q_A)} v'(w_{LH} - g(e_{LH}, x_H)) + (1 - p)(\alpha A \overline{q}_A + (1 - \alpha A)q_A) v'(w_{LH} - g(e_{LH}, x_L)).
\]

Similarly, (9) and \((IC_{HL})\) give that
\[
\lambda_{HL} = \frac{p(\alpha p(1 - \overline{q}_P) + (1 - \alpha p)(1 - q_P))}{v'(w_{LL} - g(e_{LL}, x_H))} - \gamma p(\alpha A(1 - \overline{q}_A) + (1 - \alpha A)(1 - q_A)),
\]
which together with (11) results in
\[
\gamma = \frac{\alpha p(1 - \overline{q}_P) + (1 - \alpha p)(1 - q_P)}{p(\alpha A(1 - \overline{q}_A) + (1 - \alpha A)(1 - q_A)) v'(w_{LL} - g(e_{LL}, x_H)) + (1 - p)(\alpha A(1 - \overline{q}_A) + (1 - \alpha A)(1 - q_A)) v'(w_{LL} - g(e_{LL}, x_L))}.
\]

Using (17) and (19), I now have that
\[
\frac{\alpha p \overline{q}_P + (1 - \alpha p)q_P}{1 - (\alpha p \overline{q}_P + (1 - \alpha p)q_P)} = \frac{\alpha A \overline{q}_A + (1 - \alpha A)q_A}{1 - (\alpha A \overline{q}_A + (1 - \alpha A)q_A)} \frac{p v'(w_{LH} - g(e_{LH}, x_H)) + (1 - p)v'(w_{LH} - g(e_{LH}, x_L))}{p v'(w_{LL} - g(e_{LL}, x_H)) + (1 - p)v'(w_{LL} - g(e_{LL}, x_L))}.
\]

Also, \((PC)\), \((IC_{HH})\), and \((IC_{HL})\) imply that
\[
0 = (\alpha A \overline{q}_A + (1 - \alpha A)q_A) \left( p v'(w_{LH} - g(e_{LH}, x_H)) + (1 - p)v'(w_{LH} - g(e_{LH}, x_L)) \right)
\]

\[
+ (\alpha A(1 - \overline{q}_A) + (1 - \alpha A)(1 - q_A)) \left( p v'(w_{LL} - g(e_{LL}, x_H)) + (1 - p)v'(w_{LL} - g(e_{LL}, x_L)) \right).
\]

From equations (14), (16), and (17) it follows that
\[
(p v'(w_{LH} - g(e_{LH}, x_H)))[\pi'(e_{LH}) - g_e(e_{LH}, x_H)]
\]

\[
+ (1 - p) v'(w_{LH} - g(e_{LH}, x_L)) [\pi'(e_{LH}) - g_e(e_{LH}, x_L)]
\]

\[
= pv'(w_{LH} - g(e_{LH}, x_L)) [g_e(e_{LH}, x_L) - g_e(e_{LH}, x_H)].
\]

Furthermore, it follows from equations (15), (18), and (19) that
\[
(p v'(w_{LL} - g(e_{LL}, x_H)))[\pi'(e_{LL}) - g_e(e_{LL}, x_H)]
\]

\[
+ (1 - p) v'(w_{LL} - g(e_{LL}, x_L)) [\pi'(e_{LL}) - g_e(e_{LL}, x_L)]
\]

\[
= pv'(w_{LL} - g(e_{LL}, x_L)) [g_e(e_{LL}, x_L) - g_e(e_{LL}, x_H)].
\]
Equations (12) and (13) pin down $e_{HH}^*$ and $e_{HL}^*$. The other variables are, however, determined by the above set of nonlinear equations, which can not be solved for general functional forms. Therefore, to show existence, consider the specific functional forms

$$v(\cdot) = \log(\cdot), \ g(e, x) = \frac{e^2}{2x^2}, \ \text{and} \ \pi(e) = 2e^{1/2}. $$

With these, (22) is

$$w_{ LH} = \frac{\frac{e_{ LH}^2}{2x_L} - \phi(e_{ LH})}{1 - \phi(e_{ LH})},$$

where

$$\phi(e_{ LH}) \equiv \frac{p\left(\frac{e_{ LH}}{x_L} - \frac{e_{ LH}}{x_H}\right) - (1 - p) \left(1 - \frac{e_{ LH}}{x_L} - \frac{e_{ LH}}{x_H}\right)}{p\left(\frac{e_{ LH}}{x_H} - \frac{e_{ LH}}{x_L}\right)},$$

and (23) is

$$w_{ LL} = \frac{\frac{e_{ LL}^2}{2x_L} - \xi(e_{ LL})}{1 - \xi(e_{ LL})},$$

where

$$\xi(e_{ LL}) \equiv \frac{p\left(\frac{e_{ LL}}{x_L} - \frac{e_{ LL}}{x_H}\right) - (1 - p) \left(1 - \frac{e_{ LL}}{x_L} - \frac{e_{ LL}}{x_H}\right)}{p\left(\frac{e_{ LL}}{x_H} - \frac{e_{ LL}}{x_L}\right)}.$$

Furthermore, (20) implies that

$$(\alpha_P q_P + (1 - \alpha_P) q_{ P})(1 - (\alpha_A \bar{q}_A + (1 - \alpha_A) q_{ A}) \left(p \xi(e_{ LL}) + 1 - p \right) \phi(e_{ LH})(1 - \xi(e_{ LL})) \left(1 - \frac{e_{ LH}}{x_L} - \frac{e_{ LH}}{x_H}\right) \right),$$

or

$$= \left(1 - (\alpha_P q_P + (1 - \alpha_P) q_{ P})\right)(\alpha_A \bar{q}_A + (1 - \alpha_A) q_{ A}) \left(p \phi(e_{ LH}) + 1 - p \right) \xi(e_{ LL})(1 - \phi(e_{ LH}))(1 - \frac{e_{ LH}}{x_L} - \frac{e_{ LH}}{x_H})\right),$$

Finally, (21) implies that

$$(\alpha_A \bar{q}_A + (1 - \alpha_A) q_{ A}) \log \left(\frac{e_{ LH}^2}{2x_L} - \frac{e_{ LH}}{x_L}\right) + \left(1 - (\alpha_A \bar{q}_A + (1 - \alpha_A) q_{ A}) \right) \log \left(\frac{e_{ LH}^2}{2x_L} - \frac{e_{ LH}}{x_H}\right)$$

$$- (\alpha_A \bar{q}_A + (1 - \alpha_A) q_{ A}) \log \left(e_{ LH}^2 - \frac{e_{ LH}}{x_L}\right) + p(\alpha_A \bar{q}_A + (1 - \alpha_A) q_{ A}) \log \left(p(e_{ LH}^2 - \frac{e_{ LH}}{x_L})\right)$$

$$- \left(1 - (\alpha_A \bar{q}_A + (1 - \alpha_A) q_{ A}) \right) \log \left(e_{ LL}^2 - \frac{e_{ LL}}{x_L}\right) + p \left(1 - (\alpha_A \bar{q}_A + (1 - \alpha_A) q_{ A}) \right) \log \left(p(e_{ LL}^2 - \frac{e_{ LL}}{x_L})\right)$$

$$+(1 - p)(\alpha_A \bar{q}_A + (1 - \alpha_A) q_{ A}) \log \left(p\left(\frac{e_{ LH}}{x_L} - \frac{e_{ LH}}{x_H}\right) - (1 - p)\left(e_{ LH}^2 - \frac{e_{ LH}}{x_L}\right)\right)$$

$$+(1 - p)(1 - (\alpha_P q_P + (1 - \alpha_P) q_{ P}) \log \left(p\left(\frac{e_{ LL}}{x_L} - \frac{e_{ LL}}{x_H}\right) - (1 - p)\left(e_{ LL}^2 - \frac{e_{ LL}}{x_L}\right)\right) = 0$$

14
I solve numerically for the case where $x_L = 1$, $x_H = 8$, $p = 0.5$, $Q = [0.05, 0.95]$, and a grid of points in $(\alpha_A, \alpha_P)$-space that are 0.01 apart. The parties’ utilities are then computed for each point on the grid and the principal’s utility of the bait contract is compared to her utility of a contract that does not condition on the signal. The result is that bait contracts are optimal when $\alpha_A + \alpha_P > 1$. That is, joint optimism needs to be sufficiently high.

There are two possible solutions: one has relatively high values of $w_{HH}$ and $w_{LH}$, leading to disagreement where the agent is eventually best off if the signal takes value $y_H$ and the principal is eventually best off if the signal takes value $y_L$. The other has relatively high values of $w_{HL}$ and $w_{LL}$, leading to disagreement where the agent is eventually best off if the signal takes value $y_L$ and the principal is eventually best off if the signal takes value $y_H$. Since the conditional contracts are optimal, the existence result follows.

By conditioning on the outside vague signal, the principal introduces vagueness into the otherwise precise contracting situation. That way she creates a difference in emphasis on the possible final scenarios that would otherwise not be there. If there is sufficient optimism among the parties, this is to the principal’s advantage, since it makes it easier to get the agent to participate: with a bait contract the average compensation needed for the agent’s participation is lower than with a contract that does not condition on the signal. By writing a bait contract, the principal attempts to bait the agent by offering him a very high payoff given one value of the signal. If the agent is sufficiently optimistic, he will put a lot of weight on this high payoff and take the bait. This means that the agent will accept contracts with an average compensation that would otherwise be insufficient to ensure his participation.

By bundling the contract and the side-bet on the vague signal, the principal ensures that the agent will participate in contracts with lower average compensation than he would otherwise require. Therefore, bundling dominates breaking up the problem into distinct contracting and betting problems. Breaking up the problem would require higher average compensation to the agent.

The vague environment and the assumption of OWSEU preferences together drive the results. Due to the preferences, the parties only assign weight to the best and worst lotteries in the set of possible lotteries and consequently the parties can endogenously have heterogeneous beliefs. If the environment were precise, the set of lotteries would be a singleton and
then the best and worst lotteries always coincide. In that case, the results would coincide with those of the standard model. On the other hand, when the signal is vague the best and worst lotteries can differ and then the assumption of OWSEU preferences make a difference.

The bait contracts result can explain why many publicly held companies offer stock options to rank-and-file employees. While management has a direct influence on the company as a whole, one can reasonably say that the company’s stock price is orthogonal to, e.g., the individual janitor’s performance in particular and to his employment situation more generally. Hence, this practice appears to violate Holmström’s (1979) sufficient statistic result, and the usual explanation that stock options align the incentives of the employee with those of the stock holders does not seem to apply either. The widespread use of stock options to rank-and-file employees emerged in the dot-com industry in the 1990s. The fact that it emerged in a new industry is supported by the result that the outside signal must be vague in order to optimally be conditioned upon: the future stock price of a dot-com company in the 1990s could very well be considered as a vague signal. Many dot-com companies offered stock options to compensate for lower salaries, which is again consistent with my result.

Note that the optimality of vague contracts in Theorem 1 holds even when the principal and agent have identical optimism. The important condition is that their optimism is sufficiently high (to make the optimal contract one where they disagree on which realization of the signal is best). This condition is consistent with the condition for when it is optimal to create disagreement in a model where the vague information is about the directly payoff relevant variables, see Vierø (2009b). Furthermore, quoting from the conclusion of Bergman and Jenter (2007, p. 704), their empirical evidence confirms that “firms use broad-based option compensation when boundedly rational employees are likely to be excessively optimistic about company value”. This is also in alignment with the above condition.

4 Concluding remarks

The above analysis has shown that it can be optimal for the principal to deliberately introduce vagueness into a contract by conditioning on an outside vague signal, even if this signal is uninformative about the contracting situation itself. This result can explain the use of stock options to rank-and-file employees, a practice that intuitively violates Holmström’s
sufficient statistic result.

The basic intuition and mechanism that drives the result is that vagueness gives room for the principal to affect which final scenario each of the parties put most emphasis on through the design of the contract. Hence, the vague signal gives the principal an extra instrument she can use when designing contracts. The analysis above shows that this instrument can indeed be valuable.

References


