

# Salesforce Incentives, Market Information, and Production/Inventory Planning

Fangruo Chen

Graduate School of Business, Columbia University, New York, New York 10027, fc26@columbia.edu

Salespeople are the eyes and ears of the firms they serve. They possess market knowledge that is critical for a wide range of decisions. A key question is how a firm can provide incentives to its salesforce so that it is in their interest to truthfully disclose their information about the market and to work hard. Many people have considered this question and provided solutions. Perhaps the most well-known solution is due to Gonik (1978), who proposed and implemented a clever scheme designed to elicit market information and encourage hard work. The purpose of this paper is to study Gonik's scheme and compare it with a menu of linear contracts—a solution often used in the agency literature—in a model where the market information possessed by the salesforce is important for the firm's production and inventory-planning decisions.

*Key words:* salesforce; incentives; asymmetric information; screening; production planning

*History:* Accepted by Stefanos A. Zenios, special issue editor; received December 31, 2000. This paper was with the author 7 months for 2 revisions.

## 1. Introduction

Firms rely on their salespeople to stay in touch with customers. Good salespeople know what customers need and want and the sales prospects of the market they serve. This market knowledge is critical for a wide range of decisions, such as salesforce compensation, new product development, and production-inventory planning. So, how can a firm provide incentives to its salesforce so that it is in their interest to truthfully reveal what they know about the market and, at the same time, to work hard? We address this question in a specific context where the market information is useful in making sound salesforce compensation and production/inventory decisions.

Inducing salespeople to disclose what they know about the market and providing incentives for them to work hard can sometimes be conflicting goals. To encourage the salespeople to work hard, the firm must reward them for their selling effort, which is not always proportional to the absolute sales volume. For example, when times are bad, a high level of selling effort may only generate a relatively low volume of sales. Therefore, it is sensible to set a quota that accurately reflects the potential of a sales territory and measure a salesperson's performance based on the percentage of the quota achieved. However, only the salesperson working in the territory really knows its potential. If the firm simply asks the salesperson to forecast the sales volume for his/her territory, the salesperson will try to turn in the lowest possible forecast. In other words, the firm's desire to

motivate hard work leads the salespeople to "work around the system," i.e., forecast poorly.

Many people have recognized the above problem and proposed various remedies. Gonik (1978) reported a clever scheme under which it is in the salespeople's interest to forecast accurately and to work hard. Under his scheme, the firm asks each salesperson to provide a forecast of the sales volume in his or her sales territory, and the salesperson's compensation is determined by the realized sales volume and the initial forecast. For any submitted forecast, the salesperson's compensation is a nonlinear (more precisely, piecewise linear) function of the realized sales. According to Gonik, this compensation scheme has worked successfully in IBM's Brazilian operations.<sup>1</sup>

In agency theory, the above problem is said to combine moral hazard (selling effort not observable to the firm) with adverse selection (the salesperson has superior information about the market prior to contracting with the firm). A typical solution to this type of problem is to offer a menu of contracts to the salesperson (see, e.g., Kreps 1990). By observing which contract the salesperson chooses, the firm (or principal) learns something about the market. A menu of linear contracts is often suggested as a solution and is sometimes shown to be optimal (see, e.g., Laffont and Tirole 1986, Gibbons 1987). The main purpose of this paper is to compare Gonik's solution with a menu of linear contracts.

<sup>1</sup> Variations of the Gonik story are often taught in accounting and marketing courses in business schools.

We consider a model in which a firm sells a single product through a single sales agent. The market demand for the product depends on three factors: the market condition, the selling effort by the agent, and a random noise. The agent possesses private information about the market condition, and the selling effort is not observable to the firm. The firm must design a compensation scheme for the agent and make production decisions for the product. Therefore, the agent's private information about the market condition, if revealed, benefits the firm in two ways: It increases the accuracy of the firm's measurement of the agent's selling effort, and it reduces the demand uncertainty, enabling the firm to better match supply with demand. The question is how the firm can design a contract so as to maximize its expected profit, anticipating that a production decision has to be made after contract signing, but before demand realization.

After describing the model in detail, the paper presents two benchmarks: an upper and a lower bound on the firm's achievable expected profit. The upper bound is the first-best solution: the firm's maximum expected profit if the firm could observe both the market condition and the agent's selling effort. The lower bound is obtained if the firm adopts a single linear contract, thus forgoing the opportunity to learn about the market condition. Both of these benchmarks can be obtained in closed form.

This paper then proceeds to consider menu contracts. For a menu of linear contracts, the optimal values of the contract parameters can be obtained in closed form, leading to a simple calculation of the firm's optimal expected profit. For Gonik's (1978) menu of nonlinear contracts (taking his contract structure while ignoring his particular choices of parameter values), the problem becomes much harder. However, based on several analytical results, we are able to devise an efficient algorithm for computing the optimal values of the contract parameters and the corresponding expected profit for the firm. We then compare these two solutions in a numerical experiment. Our results show that the Gonik solution is dominated by a menu of linear contracts. This appears to be due to the Gonik solution's inability to induce different effort levels for different market conditions. The numerical analysis also compares these solutions with the above benchmarks and shows how the differences among these scenarios vary as some of the model parameters change.

The salesforce compensation problem has been widely studied in the marketing literature; see Coughlan (1993) for a comprehensive review. Many of the existing models assume, as is done here, that the sales response function has a random noise term; i.e., the total sales generated by a given level of selling effort is random. This, together with the fact that

the selling effort is often unobservable to the firm, leads to the moral hazard problem: the problem of motivating the salespeople to work, given that their reward can only be based on an imperfect signal of effort. If, in addition, the salespeople have superior information about the sales response function (the productivity of selling effort, the sensitivity of customers to price changes, the sales prospects, etc.), then the firm is at an informational disadvantage in terms of the sales environment. This is the adverse-selection problem mentioned earlier. The basic model of the moral hazard problem has been studied and refined by many in the economics/agency-theory literature; see, e.g., Shavell (1979); Harris and Raviv (1978, 1979); Holmstrom (1979, 1982); and Grossman and Hart (1983). This machinery was then introduced to the salesforce compensation literature in marketing by Basu et al. (1985). Many marketing researchers have since extended the basic moral hazard model (one product, one salesperson, symmetric information) to models with asymmetric information and multiple sales territories; see, e.g., Lal (1986), Lal and Staelin (1986), Rao (1990), and Raju and Srinivasan (1996). Our model is a combination of moral hazard and adverse selection. Closely related to the current paper is Mantrala and Raman (1990), who have formalized Gonik's idea and focused on analyzing the sales agent's response to a Gonik scheme. Based on this analysis, they suggested several guidelines as to how the firm can go about choosing the contract parameters under Gonik's scheme. One of our contributions here is to directly address the problem facing the firm (the agent's response is only a subproblem) and to provide analytical results that lead to an efficient algorithm for computing the optimal contract parameters. It is also worth noting that a novel feature of our model as compared to Mantrala and Raman (1990), and for that matter, most models in agency theory, is that the principal (the firm) has an activity—i.e., production/inventory planning—that stands to benefit from the revelation of the agent's private information. On the other hand, the operations literature often takes as given the knowledge about demand (e.g., the characterization of the demand process), whereas here part of the demand information is unknown to the firm and salesforce incentives are required for its revelation. We therefore contribute to the growing body of research in an area often labeled the "marketing-operations interface." For some of this interface research with a specific connection to salesforce incentives, see Dearden and Lilien (1990), Porteus and Whang (1991), and Chen (2000).

The rest of this paper is organized as follows. Section 2 describes the model and presents some preliminary results. Section 3 presents two benchmark solutions: the first-best solution and one based on a

single linear contract. Section 4 shows how an optimal menu of linear contracts can be determined in closed form. Section 5 deals with the Gonik solution. Section 6 presents the numerical results. Concluding remarks are in §7.

## 2. The Model

A firm sells a single product through a sales agent. The total sales or demand,  $X$ , is determined by the agent's selling effort ( $a$ ), the market condition ( $\theta$ ), and a random noise ( $\epsilon$ ) via the following additive form:

$$X = a + \theta + \epsilon,$$

where  $a$  is a nonnegative real number, and  $\theta$  and  $\epsilon$  are independent random variables with  $\Pr(\theta = \theta_H) = \rho$  and  $\Pr(\theta = \theta_L) = 1 - \rho$  for  $0 < \rho < 1$  and  $\theta_H > \theta_L > 0$ , and  $\epsilon \sim N(0, \sigma^2)$ . Moreover, assume that  $\theta_L$  is sufficiently large that the probability of  $X$  being negative is negligible.

The relationship between the firm and the salesperson is that of a principal and an agent in the sense that the latter sells the product on behalf of the former. The principal designs the agent's wage contract and makes production decisions, while the agent—endowed with private information about the market condition—decides whether or not to accept a contract and, if so, how much selling effort to exert. More specifically, we assume the following sequence of events: (1) The firm (or principal) offers a menu of wage contracts; (2) the agent privately observes the value of  $\theta$ ; (3) the agent decides whether or not to participate (work for the firm) and if so, which contract to sign; (4) under a signed contract, the firm determines the production quantity, and the agent makes the effort decision; and (5) both parties observe the total sales (i.e., the value of  $X$ ). The firm cannot directly observe the agent's effort level, and thus must compensate the agent based on the realized value of  $X$ .

The problem facing the firm is thus a mixture of moral hazard (postcontractual opportunism associated with the effort decision) and adverse selection (precontractual asymmetric information regarding the market condition). A typical response to this type of problem is to offer a menu of contracts to the agent, who—knowing the market condition—then chooses one of them to sign. By observing the choice made by the agent, the firm may learn something about the market condition (this is often referred to as screening in economics), and this knowledge may be helpful in making production decisions.

Consider the agent's decisions when offered a menu of contracts. First, he would consider each contract on the menu and determine the maximum expected utility that could be obtained under the contract. Suppose  $s(\cdot)$  is the contract being considered.

(Thus,  $s(x)$  is the wage paid to the agent if the total sales is  $x$ .) Assume the agent's utility for net income  $z$  is  $U(z) = -e^{-rz}$  with  $r > 0$ . (The negative exponential utility function is widely used in the agency literature.) Note that  $U(\cdot)$  is increasing and concave, implying that the agent is risk averse. The net income is the wage received,  $s(X)$ , minus the cost of effort, which is assumed to be  $V(a) = a^2/2$ , an increasing, convex function.<sup>2</sup> To determine the maximum expected utility achievable under  $s(\cdot)$ , the agent solves the following optimization problem:

$$\max_a E[-e^{-r(s(X)-V(a))}].$$

Recall that the agent has already observed the value of  $\theta$  when evaluating the contract. Therefore, the above expectation is with respect to  $\epsilon$ , given the observed value of  $\theta$ . For convenience, we say that the agent is of high type if he has observed  $\theta = \theta_H$ , and low type otherwise. The optimal effort decision thus depends not only on the contract  $s(\cdot)$ , but also the agent's type. Let  $a(s, t)$  be the optimal effort decision given contract  $s(\cdot)$  and agent's type  $t$  ( $= H$  or  $L$ ). Let  $u(s, t)$  be the corresponding expected utility for the agent, i.e., the maximum achievable expected utility under  $s$  and  $t$ . If  $u(s, t)$  is greater than or equal to  $-U_0$ —the agent's reservation utility representing the best outside opportunity for the agent—then  $s$  is said to be acceptable to the agent.<sup>3</sup> Among all the contracts on the menu, the agent chooses the one with the highest achievable expected utility and participates if this utility level exceeds  $-U_0$ . Throughout this paper, we will assume that the menu offered by the principal contains at least one contract that is acceptable to the high type and, similarly, at least one acceptable to the low type. This avoids the situation where the principal is left without a salesperson.

We now turn to the principal's problem. The above sequence of events implies that the firm must make its production decision before observing the total sales. This is reasonable when the customers demand fast delivery of their orders and the production lead time is relatively long. (It is thus impossible to follow make-to-order.) Let  $Q$  be the production quantity. Let  $c$  be the cost per unit produced. When supply does not match demand, additional costs are incurred. If  $X \leq Q$ , the excess supply is salvaged at  $p$  per unit. On the other hand, if  $X > Q$ , the excess demand must be satisfied via an emergency action such as a special production run at a cost of  $c'$  per unit. Let the unit

<sup>2</sup> Other functional forms for the cost of effort can be used without fundamentally changing the analysis.

<sup>3</sup> It is reasonable to assume that the reservation utility does not depend on the agent's type, because what distinguishes the high type from the low type is the market condition, something unrelated to the agent's intrinsic quality.

selling price be  $1 + c$  (the profit margin is thus normalized to 1). To avoid trivial cases, assume  $p < c < c' < 1 + c$ .<sup>4</sup> The firm makes contract as well as production decisions with the objective of maximizing its expected profit (the principal is thus risk neutral). If  $s(\cdot)$  is the contract signed by the agent, the firm's profit is

$$\begin{aligned} & (1 + c)X - s(X) - cQ + p(Q - X)^+ - c'(Q - X)^- \\ & = X - s(X) \\ & \quad - [(c - p)(Q - X)^+ + (c' - c)(Q - X)^-], \end{aligned} \quad (1)$$

where  $w^+ = \max\{w, 0\}$  and  $w^- = \max\{-w, 0\}$ . Note that the optimal production quantity minimizes

$$E[(c - p)(Q - X)^+ + (c' - c)(Q - X)^-], \quad (2)$$

where the expectation is with respect to  $X$  given the principal's knowledge about the market condition and the agent's selling effort (inferred, not observed) after a contract is signed.

Because there are only two possible market conditions (i.e., the agent has only two possible types), the firm needs to offer at most two contracts. Therefore, there are only two possibilities: Either the two agent types choose the same contract (in this case the principal is effectively offering one contract), or they choose different contracts. If the former, we say that the two types are "pooled," and if the latter, "separated." We next formulate the principal's problem under each of these scenarios.

Consider the pooling case. Let  $s(\cdot)$  be the contract offered and signed. This contract induces different levels of selling effort depending on the agent's type: The high type exerts  $a_H = a(s, H)$ , and the low type  $a_L = a(s, L)$ . (The principal does not observe the amount of effort, but she can correctly anticipate it for each type.) As mentioned earlier, the contract must be acceptable to both types; i.e.,  $u(s, H) \geq -U_0$  and  $u(s, L) \geq -U_0$ . Because only one contract is offered, the principal does not learn any new information with regard to the market condition after it is signed. Therefore, before making the production decision, the firm's knowledge about the distribution of  $X$  is: With probability  $\rho$ , the agent is of high type (i.e.,  $\theta = \theta_H$ ), he will exert effort  $a_H$ , and thus  $X \sim N(a_H + \theta_H, \sigma^2)$ ; with probability  $1 - \rho$ , the agent is of low type (i.e.,  $\theta = \theta_L$ ), he will exert effort  $a_L$ , and thus  $X \sim N(a_L + \theta_L, \sigma^2)$ . In other words,  $X$  is a mixture of two normal random

variables with known means and variances. With this distribution, one can minimize (2) over  $Q$  to identify the optimal production quantity. This is a standard newsvendor problem. Let  $Q_0^*$  be the optimal quantity, and  $G_0^*$  the corresponding value of (2). The problem of finding an optimal (pooling) contract can be stated as

$$\begin{aligned} & \max_{s, a_H, a_L} \rho E[X - s(X) | a = a_H, \theta = \theta_H] \\ & \quad + (1 - \rho) E[X - s(X) | a = a_L, \theta = \theta_L] - G_0^* \\ & \text{s.t. } a_H = a(s, H) \quad (\text{IC-H}), \\ & \quad a_L = a(s, L) \quad (\text{IC-L}), \\ & \quad u(s, H) \geq -U_0 \quad (\text{IR-H}), \\ & \quad u(s, L) \geq -U_0 \quad (\text{IR-L}), \end{aligned}$$

where the first two constraints are incentive compatibility (IC) constraints indicating that the effort level for each agent type is indeed optimal for the agent (thus compatible with the agent's objective), and the remaining two constraints indicate that the agent, regardless of his type, is better off participating relative to his outside opportunities, i.e., individual rationality (IR).

Now consider the separating case where the two agent types choose different contracts. Let  $s_H(\cdot)$  be the contract chosen by the high type, and  $s_L(\cdot)$  chosen by the low type. Again, the principal can anticipate the amount of selling effort under each type:  $a_H = a(s_H, H)$  for the high type, and  $a_L = a(s_L, L)$  for the low type. Unlike the pooling case, here the principal discovers the realized value of  $\theta$  by observing the contract choice made by the agent. If the agent chooses the contract that is preferred by the high type (i.e.,  $s_H(\cdot)$ ), then  $\theta = \theta_H$ . Similarly, if  $s_L(\cdot)$  is chosen,  $\theta = \theta_L$ . Consequently, the principal can make more accurate production decisions. If  $s_H(\cdot)$  is chosen, then  $X \sim N(a_H + \theta_H, \sigma^2)$ , and thus the optimal production quantity is  $Q_H^* = a_H + \theta_H + q^*$ , where  $q^*$  is the minimizer of  $g(q) = E[(c - p)(q - \epsilon)^+ + (c' - c)(q - \epsilon)^-]$ . Let  $G_1^* = g(q^*)$ , the minimum expected production/inventory costs if  $s_H(\cdot)$  is chosen. On the other hand, if  $s_L(\cdot)$  is chosen, then  $X \sim N(a_L + \theta_L, \sigma^2)$ , and the optimal production quantity is thus  $Q_L^* = a_L + \theta_L + q^*$ . In this case, the minimum expected production/inventory costs are again  $G_1^*$ . An optimal menu of contracts can be found by solving the following problem:

$$\begin{aligned} & \max_{s_H(\cdot), s_L(\cdot), a_H, a_L} \rho E[X - s_H(X) | a = a_H, \theta = \theta_H] \\ & \quad + (1 - \rho) E[X - s_L(X) | a = a_L, \theta = \theta_L] - G_1^* \\ & \text{s.t. } a_H = a(s_H, H) \quad (\text{IC-H}), \\ & \quad a_L = a(s_L, L) \quad (\text{IC-L}), \\ & \quad u(s_H, H) \geq u(s_L, H) \quad (\text{IC-HL}), \\ & \quad u(s_L, L) \geq u(s_H, L) \quad (\text{IC-LH}), \end{aligned}$$

<sup>4</sup> The case where the excess demand is lost represents an interesting question. If we assume that  $X$  is still observed even if it exceeds  $Q$  so that compensation can be written as  $s(X)$ , then essentially the same analysis applies; simply replace  $c'$  with  $1 + c$ . On the other hand, if  $X$  is not observable beyond  $Q$ , what should compensation be based on? This is an interesting question for future research.

$$u(s_H, H) \geq -U_0 \quad (\text{IR-H}),$$

$$u(s_L, L) \geq -U_0 \quad (\text{IR-L}),$$

where the first two constraints are the incentive compatibility constraints for the agent's effort choices, the third and fourth constraints ensure that the agent picks the contract that is intended for his type, and the last two are the individual rationality constraints. Implicit in the above formulation is that the two contracts,  $s_H(\cdot)$  and  $s_L(\cdot)$ , are distinct. Otherwise,  $G_1^*$  must be replaced with  $G_0^*$ , and the formulation reduces to the one under pooling. Also, if the agent is indifferent between two contracts, we assume he chooses the one intended for his type.

The following useful result is well known. When the agent's net income is normally distributed and his utility function has the negative-exponential form, his expected utility has a simple expression. Let  $Y$  be the normally distributed net income. The agent's expected utility is thus  $E[-e^{-rY}]$ . The certainty equivalent of  $Y$ , denoted by  $CE[Y]$ , is the fixed net income that provides the agent with a utility level equal to  $E[-e^{-rY}]$ ; i.e.,  $-e^{-rCE[Y]} = E[-e^{-rY}]$ . It can be easily verified that

$$CE[Y] = \mu_Y - \frac{1}{2}r\sigma_Y^2,$$

where  $\mu_Y$  is the mean of  $Y$  and  $\sigma_Y^2$  the variance. Maximizing the expected utility is equivalent to maximizing the certainty equivalent.

### 3. Two Benchmarks

This section presents two benchmarks for the firm's performance. One assumes that the firm is omniscient (able to observe both the market condition and the agent's effort), while the other assumes that the firm offers a single contract, thus pooling together the two agent types and forgoing the opportunity to discover the market condition. These benchmarks provide upper and lower bounds on the firm's expected profit.

#### 3.1. The First-Best Solution

To obtain the first-best solution, assume that the firm is able to see the market condition, i.e., the value of  $\theta$ , and to observe the amount of selling effort exerted by the agent. In this scenario, the firm can specify the effort level to be exerted by the agent under each possible market condition and pay the agent a fixed minimum wage that satisfies the participation constraint. Paying the agent a fixed wage is optimal for the firm because of the difference in risk attitude: The agent is risk averse, but the principal is risk neutral. The firm's problem is, therefore, to identify an effort level and a fixed wage that are acceptable to the agent and at the same time maximize the firm's expected profit.

Under fixed wage  $w$ , the participation constraint becomes  $-e^{-r(w-V(a))} \geq -U_0$ , or

$$w \geq V(a) - \frac{\ln U_0}{r} \stackrel{\text{def}}{=} w_0(a),$$

where  $w_0(a)$  is the minimum wage for given effort level  $a$ .

Suppose  $\theta = \theta_H$ . Replacing  $s(X)$  with  $w_0(a)$  in (1), we have the firm's expected profit as a function of  $a$ :

$$\begin{aligned} & E \left\{ X - \left( V(a) - \frac{\ln U_0}{r} \right) \right. \\ & \quad \left. - [(c-p)(Q-X)^+ + (c'-c)(Q-X)^-] \right\} \\ & = \theta_H + a - V(a) + \frac{\ln U_0}{r} \\ & \quad - E[(c-p)(Q - \theta_H - a - \epsilon)^+ \\ & \quad \quad + (c'-c)(Q - \theta_H - a - \epsilon)^-] \\ & = \theta_H + a - V(a) + \frac{\ln U_0}{r} - g(Q - \theta_H - a), \end{aligned}$$

where the first equality follows from  $X = \theta_H + a + \epsilon$  and  $E[\epsilon] = 0$ . For any given value of  $a$ , the optimal production quantity is  $Q^* = \theta_H + a + q^*$ , with the minimum production/inventory costs equal to  $g(Q^* - \theta_H - a) = g(q^*) = G_1^*$ , a value independent of  $a$ . Therefore, the objective function reduces to

$$\theta_H + a - V(a) + \frac{\ln U_0}{r} - G_1^*,$$

a concave function in  $a$  with the first-order condition  $1 - V'(a) = 0$ . Because  $V(a) = a^2/2$ , the optimal effort level is  $a^* = 1$ . The firm's expected profit is thus  $\theta_H + 0.5 + (\ln U_0)/r - G_1^*$ .

Similarly, if  $\theta = \theta_L$ , the optimal effort level is  $a^* = 1$  and the optimal production quantity is  $Q^* = \theta_L + 1 + q^*$ , giving the firm an expected profit of  $\theta_L + 0.5 + (\ln U_0)/r - G_1^*$ .

Combining the above two cases, the firm's ex ante (before observing the market condition) expected profit is  $\bar{\theta} + 0.5 + (\ln U_0)/r - G_1^*$ , where  $\bar{\theta} = \rho\theta_H + (1-\rho)\theta_L$ , the expected market condition.

#### 3.2. A Single Linear Contract

Suppose the firm offers a single wage contract that is linear in total sales, i.e.,  $s(x) = \alpha x + \beta$ , where  $\beta$  is the fixed salary and  $\alpha \geq 0$  is the commission rate. This is the pooling scenario discussed in the previous section, where a formulation for the firm's problem is also presented.

Suppose the agent is of type  $\theta$ , and consider his effort decision under the linear contract. Note that the agent's net income is

$$\begin{aligned} s(X) - V(a) & = \alpha(\theta + a + \epsilon) + \beta - V(a) \\ & \sim N(\alpha(\theta + a) + \beta - V(a), \alpha^2\sigma^2), \end{aligned}$$

with its certainty equivalent

$$\alpha(\theta + a) + \beta - V(a) - \frac{1}{2}r\alpha^2\sigma^2.$$

Maximizing the certainty equivalent with respect to  $a$ , the agent solves the first-order condition,  $\alpha - V'(a) = 0$ , yielding the optimal effort level  $a^* = \alpha$ . Note that this effort decision is independent of the agent's type  $\theta$ . With the optimal effort, the agent's certainty equivalent becomes

$$\alpha\theta + \beta + \frac{\alpha^2}{2}(1 - r\sigma^2). \quad (3)$$

Now consider the participation constraints. Note that the certainty equivalent of the agent's outside opportunities is  $-(\ln U_0)/r$ . From (3), the participation constraints are

$$\alpha\theta + \beta + \frac{\alpha^2}{2}(1 - r\sigma^2) \geq -\frac{\ln U_0}{r}, \quad \theta = \theta_H, \theta_L.$$

Because  $\alpha \geq 0$ , we only need to consider the constraint associated with the low type.

Recall from the previous section that the firm's objective function under the pooling case is

$$\begin{aligned} & \rho E[X - s(X) | a = a_H, \theta = \theta_H] \\ & + (1 - \rho)E[X - s(X) | a = a_L, \theta = \theta_L] - G_0^*. \end{aligned}$$

Using  $a_H = a_L = \alpha$  and simplifying, we have the firm's objective function

$$(1 - \alpha)(\bar{\theta} + \alpha) - \beta - G_0^*.$$

The firm's problem is thus

$$\begin{aligned} & \max_{\alpha, \beta} (1 - \alpha)(\bar{\theta} + \alpha) - \beta - G_0^* \\ & \text{s.t. } \alpha\theta_L + \beta + \frac{\alpha^2}{2}(1 - r\sigma^2) \geq -\frac{\ln U_0}{r}. \end{aligned}$$

Note that the optimal solution must make the constraint binding; i.e.,

$$\beta = -\alpha\theta_L - \frac{\alpha^2}{2}(1 - r\sigma^2) - \frac{\ln U_0}{r}.$$

Substituting this expression for  $\beta$  in the firm's objective function leads to a concave function of  $\alpha$ . Maximizing it over  $\alpha \geq 0$ , we obtain

$$\alpha^* = \frac{1}{1 + r\sigma^2} \max\{1 - \rho(\theta_H - \theta_L), 0\}.$$

It is interesting to note that when  $\theta_H - \theta_L$  is sufficiently large, the optimal solution is to offer the agent a fixed salary (with no incentive pay). Also note that if  $\theta_H - \theta_L = 0$ —i.e., there is no uncertainty about the market condition—the firm always offers a pay package with a positive commission rate. This

suggests that one reason for the prevalence of fixed salaries in practice is the firm's uncertainty about the market condition (and its desire to offer a single contract). The intuition for this result is that if  $\alpha$  is positive, the high-type agent obtains a surplus because of the low-type agent's participation constraint, and this surplus is proportional to  $\theta_H - \theta_L$ . When  $\theta_H - \theta_L$  is very large, the firm is paying the agent too much; i.e., the high type's surplus exceeds the incremental revenue generated by the agent's effort (which is equal to  $\alpha$ ). In this case, the firm is better off with a fixed salary, which can be chosen to make the agent's participation constraint binding for both types.

#### 4. A Menu of Linear Contracts

This section takes up the case where the firm offers two linear contracts to the agent, who then chooses one based on his knowledge of the market condition. By observing which of the two contracts the agent chooses, the firm discovers the agent's type (i.e., market condition). This information enables the firm to tailor its production decision to the market condition, reducing production/inventory costs. In short, we consider the formulation presented in §2 for the separating case by restricting to linear contracts.

Recall from §3.2 that the optimal response of the type- $\theta$  agent to a linear contract  $s(x) = \alpha x + \beta$  is to exert effort  $a^* = \alpha$ , yielding a maximum expected utility with certainty equivalent  $\alpha\theta + \beta + ((1 - r\sigma^2)/2)\alpha^2$ . By offering  $s(\cdot)$ , the principal's expected profit (excluding the production/inventory costs) is  $E[X - s(X)] = (1 - \alpha)(\alpha + \theta) - \beta$  if the market condition is  $\theta$ . These results are useful in formulating the firm's problem under a menu of linear contracts.

Consider the separating scenario. Let  $s_H(x) = \alpha_H x + \beta_H$  be the contract intended for the high-type agent, and  $s_L(x) = \alpha_L x + \beta_L$  the contract intended for the low-type agent, with  $\alpha_H, \alpha_L \geq 0$ . The firm's optimization problem can be written as

$$\begin{aligned} & \max_{\alpha_H, \beta_H, \alpha_L, \beta_L} \rho[(1 - \alpha_H)(\alpha_H + \theta_H) - \beta_H] \\ & \quad + (1 - \rho)[(1 - \alpha_L)(\alpha_L + \theta_L) - \beta_L] - G_1^* \\ & \text{s.t. } \alpha_H\theta_H + \beta_H + \frac{1 - r\sigma^2}{2}\alpha_H^2 \\ & \quad \geq \alpha_L\theta_H + \beta_L + \frac{1 - r\sigma^2}{2}\alpha_L^2 \quad (\text{IC-HL}), \\ & \quad \alpha_L\theta_L + \beta_L + \frac{1 - r\sigma^2}{2}\alpha_L^2 \\ & \quad \geq \alpha_H\theta_L + \beta_H + \frac{1 - r\sigma^2}{2}\alpha_H^2 \quad (\text{IC-LH}), \\ & \quad \alpha_H\theta_H + \beta_H + \frac{1 - r\sigma^2}{2}\alpha_H^2 \\ & \quad \geq -\frac{\ln U_0}{r} \quad (\text{IR-H}), \end{aligned}$$

$$\begin{aligned} & \alpha_L \theta_L + \beta_L + \frac{1-r\sigma^2}{2} \alpha_L^2 \\ & \geq -\frac{\ln U_0}{r} \quad (\text{IR-L}). \end{aligned}$$

The following observations significantly simplify the above optimization problem. First, the constraints (IC-HL) and (IR-L) together imply (IR-H) because  $\theta_H > \theta_L$  and  $\alpha_L \geq 0$ . Thus, eliminate (IR-H). Second, in the reduced problem, observe that (IC-HL) must hold as an equality, because otherwise one can always decrease  $\beta_H$  without violating the constraints, and improve the principal's payoff. Thus, change the inequality in (IC-HL) to equality. Third, (IC-LH) can be replaced with  $\alpha_H \geq \alpha_L$ . To see this, note that (IC-LH)  $\Leftrightarrow$  (IC-HL) with equality + (IC-LH), i.e.

$$\begin{aligned} & \left( \alpha_H \theta_H + \beta_H + \frac{1-r\sigma^2}{2} \alpha_H^2 \right) + \left( \alpha_L \theta_L + \beta_L + \frac{1-r\sigma^2}{2} \alpha_L^2 \right) \\ & \geq \left( \alpha_L \theta_H + \beta_L + \frac{1-r\sigma^2}{2} \alpha_L^2 \right) \\ & \quad + \left( \alpha_H \theta_L + \beta_H + \frac{1-r\sigma^2}{2} \alpha_H^2 \right), \end{aligned}$$

which can be simplified to  $\alpha_H \geq \alpha_L$  because  $\theta_H > \theta_L$ . Now replace (IC-LH) with  $\alpha_H \geq \alpha_L$ . Finally, (IR-L) must hold as an equality because otherwise one can decrease  $\beta_H$  and  $\beta_L$  by the same amount without violating any constraints and increase the objective function. Changing (IR-L) to an equality and applying all of the above changes, we have the following equivalent problem:

$$\begin{aligned} & \max_{\alpha_H, \beta_H, \alpha_L, \beta_L} \quad \rho[(1-\alpha_H)(\alpha_H + \theta_H) - \beta_H] \\ & \quad + (1-\rho)[(1-\alpha_L)(\alpha_L + \theta_L) - \beta_L] - G_1^* \\ \text{s.t.} \quad & \alpha_H \theta_H + \beta_H + \frac{1-r\sigma^2}{2} \alpha_H^2 \\ & = \alpha_L \theta_H + \beta_L + \frac{1-r\sigma^2}{2} \alpha_L^2 \quad (\text{IC-HL}), \\ & \alpha_L \theta_L + \beta_L + \frac{1-r\sigma^2}{2} \alpha_L^2 = -\frac{\ln U_0}{r} \quad (\text{IR-L}), \\ & \alpha_H \geq \alpha_L. \end{aligned}$$

Therefore, in the optimal solution the high-type agent is indifferent between the two contracts (see (IC-HL)), and the low-type agent is indifferent between  $s_L(\cdot)$  and his outside opportunities.

We are now ready to solve the principal's problem. First, use (IR-L) to solve for  $\beta_L$ , and use this result in (IC-HL) to solve for  $\beta_H$ . That is,

$$\beta_L = -\frac{\ln U_0}{r} - \alpha_L \theta_L - \frac{1-r\sigma^2}{2} \alpha_L^2$$

and

$$\beta_H = -\frac{\ln U_0}{r} + \alpha_L(\theta_H - \theta_L) - \alpha_H \theta_H - \frac{1-r\sigma^2}{2} \alpha_H^2.$$

Using these expressions to replace  $\beta_H$  and  $\beta_L$  in the objective function, we have the following:

$$\begin{aligned} & \max_{\alpha_H, \alpha_L} \quad \rho \left[ \alpha_H + \theta_H - \alpha_L(\theta_H - \theta_L) - \frac{1+r\sigma^2}{2} \alpha_H^2 \right] \\ & \quad + (1-\rho) \left[ \alpha_L + \theta_L - \frac{1+r\sigma^2}{2} \alpha_L^2 \right] + \frac{\ln U_0}{r} - G_1^* \\ \text{s.t.} \quad & \alpha_H \geq \alpha_L \geq 0, \end{aligned}$$

where the nonnegative constraints on  $\alpha_H$  and  $\alpha_L$  were mentioned earlier, but are now explicitly included in the formulation. Note that the objective function is separable and concave in  $\alpha_H$  and  $\alpha_L$ . It is then straightforward to obtain the optimal solution, which is

$$\begin{aligned} \alpha_H^* &= \frac{1}{1+r\sigma^2} \quad \text{and} \\ \alpha_L^* &= \frac{1}{1+r\sigma^2} \max \left\{ 1 - \frac{\rho}{1-\rho}(\theta_H - \theta_L), 0 \right\}. \end{aligned}$$

We close this section with an examination of the agent's selling effort under various scenarios. First, suppose the firm knows the market condition but cannot observe the agent's effort. In this case, the firm can have two linear contracts, one for each agent type. It is easy to verify that under the optimal linear contracts, the selling efforts for the two agent types are the same:

$$a_L^0 = a_H^0 = \frac{1}{1+r\sigma^2}.$$

This should be intuitive because the two agent types have the same productivity of effort in terms of revenue generation, and they have the same cost of effort. Second, suppose the firm does not know the market condition and cannot observe the agent's effort (this is the case we have been considering all along). Here the firm can either restrict itself to a single linear contract (as in §3.2) or offer a menu of two linear contracts (as in this section). Under a single linear contract, the selling efforts for the two agent types are still the same:

$$a_L^1 = a_H^1 = \frac{1}{1+r\sigma^2} \max\{1 - \rho(\theta_H - \theta_L), 0\}.$$

Under a menu of two linear contracts, the selling efforts are different:

$$a_L^2 = \frac{1}{1+r\sigma^2} \max \left\{ 1 - \frac{\rho}{1-\rho}(\theta_H - \theta_L), 0 \right\}, \quad a_H^2 = \frac{1}{1+r\sigma^2}.$$

Note that  $a_L^0 \geq a_L^1 \geq a_L^2$  and  $a_H^0 \geq a_H^1 \leq a_H^2 = a_H^0$ . Therefore, the firm's lack of market information generally leads to reduced selling effort, and adopting a menu of contracts rather than a single contract increases the effort of the high-type agent, but reduces the effort of the low type.

## 5. The Gonik Proposal

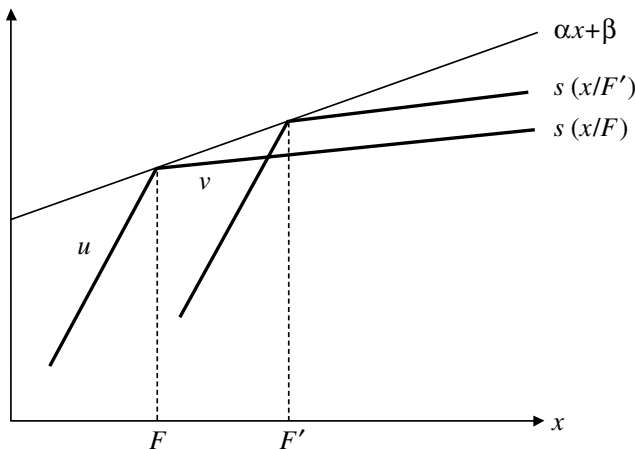
Gonik (1978) has designed a salesforce compensation scheme to motivate the agent to be precise about the total sales and, at the same time, to work hard. Under his scheme, the firm asks the salesperson to submit a forecast of the total sales. If the forecast is  $F$ , then  $s(x | F)$ —a given function of the actual total sales  $x$  parameterized by  $F$ —is the compensation for the agent. Therefore, the firm is effectively offering a menu of contracts; by submitting a forecast, the agent chooses a particular contract from the menu. The agent submits his forecast after observing the market condition. Therefore, the forecast may contain useful information about the market condition, which the firm can then use in making its production decisions. A key feature of the Gonik scheme is that if the agent expects to sell  $x$  units, it is in his best interest to submit a forecast that is equal to  $x$ ; and that given any forecast, the agent’s compensation is increasing in the actual total sales, providing the agent with incentives to generate more sales.

Let us formalize the above idea. Recall that  $s(x | F)$  is the agent’s compensation, with  $F$  being the forecast and  $x$  the (actual) total sales. To motivate the agent to provide an accurate forecast, we must have  $s(x | x) \geq s(x | F)$  for all  $F$  and  $x$ . To motivate the agent to work hard,  $s(x | F)$  must be increasing in  $x$  for any  $F$ . The following example satisfies these two conditions:  $s(x | x) = \alpha x + \beta$  for all  $x$ ; and for any  $x$  and  $F$ ,

$$s(x | F) = \begin{cases} s(F | F) - u(F - x), & x \leq F, \\ s(F | F) + v(x - F), & x > F, \end{cases} \quad (4)$$

where  $\alpha$ ,  $\beta$ ,  $u$ , and  $v$  are contract parameters chosen by the firm with  $u > \alpha > v > 0$ . Figure 1 depicts two possible contracts under the above scheme. This menu of piecewise linear (in  $x$ ) contracts—with properly chosen parameters—is precisely what Gonik (1978) proposed and implemented at IBM Brazil’s sales

Figure 1 Gonik’s Menu of Piecewise Linear Contracts



operations. The purpose for the rest of this section is to show how the optimal values of  $\alpha$ ,  $\beta$ ,  $u$ , and  $v$  can be determined to maximize the firm’s expected profit.

### 5.1. The Agent’s Problem

Consider the problem facing the agent when given a menu of contracts specified in (4). The agent makes two decisions: the forecast  $F$  and the selling effort  $a$ . These decisions are made after the agent sees the market condition ( $\theta$ ), but before the demand noise  $\epsilon$  is realized. The objective is to maximize the agent’s expected utility, which is

$$E[-e^{-r(s(X|F)-V(a))}] = -e^{rV(a)} E[e^{-rs(X|F)}].$$

Recall that  $X = \theta + a + \epsilon \sim N(\theta + a, \sigma^2)$ . Using (4), we have

$$E[e^{-rs(X|F)}] = \int_{-\infty}^F e^{-r(s(F|F)-u(F-x))} \frac{1}{\sigma} \phi\left(\frac{x-\theta-a}{\sigma}\right) dx + \int_F^{\infty} e^{-r(s(F|F)+v(x-F))} \frac{1}{\sigma} \phi\left(\frac{x-\theta-a}{\sigma}\right) dx,$$

where  $\phi(\cdot)$  is the standard normal density function. Using  $s(F | F) = \alpha F + \beta$  in the above expression and simplifying, we have

$$E[e^{-rs(X|F)}] = e^{-r(\alpha F + \beta)} \left[ e^{ruz + \frac{1}{2}(ru\sigma)^2} \Phi\left(\frac{z}{\sigma} + ru\sigma\right) + e^{rvz + \frac{1}{2}(rv\sigma)^2} \bar{\Phi}\left(\frac{z}{\sigma} + rv\sigma\right) \right], \quad (5)$$

where  $z = F - a - \theta$ ,  $\Phi$  is the standard normal cdf, and  $\bar{\Phi} = 1 - \Phi$ . Therefore, the agent’s optimization problem can be stated as

$$\max_{a, z} -e^{r(V(a)-\alpha a)} e^{-r\alpha\theta-r\beta} \psi(z), \quad (6)$$

where

$$\psi(z) = e^{r(u-\alpha)z + \frac{1}{2}(ru\sigma)^2} \Phi\left(\frac{z}{\sigma} + ru\sigma\right) + e^{r(v-\alpha)z + \frac{1}{2}(rv\sigma)^2} \bar{\Phi}\left(\frac{z}{\sigma} + rv\sigma\right),$$

which also depends on  $\alpha$ ,  $u$ , and  $v$ , but is independent of  $\beta$ .

Note that the agent’s objective function in (6) is separable in  $a$  and  $z$ . The optimal selling effort  $a^*$  minimizes  $V(a) - \alpha a$ . With  $V(a) = a^2/2$ ,  $a^* = \alpha$ . It is interesting that the optimal effort level is entirely determined by only one contract parameter,  $\alpha$ , and it is independent of the agent’s type. Now let  $\psi(z)$  be minimized at  $z^*$ . The optimal forecast decision then is,  $F^* = z^* + a^* + \theta = z^* + \alpha + \theta$ . Note that  $z^*$  depends on  $\alpha$ ,  $u$ , and  $v$ , but it is independent of  $\beta$  and the



agent's type. Therefore, the high-type agent forecasts  $F_H = z^* + \alpha + \theta_H$  and the low-type forecasts  $F_L = z^* + \alpha + \theta_L$ . As a result, after receiving the agent's forecast, the firm discovers the market condition and makes the production decision accordingly. This is the separating case discussed in §2.

We pause here to collect several results on  $\psi(\cdot)$ , which are useful in determining  $z^*$ .

**THEOREM 1.**  $\psi(\cdot)$  is quasi-convex with  $\psi(z) \geq 1$  for all  $z$ .

**PROOF.** Define

$$\begin{aligned} f_1(\alpha, z) &= e^{-r\alpha z} \quad \text{and} \\ f_2(z) &= e^{ruz + \frac{1}{2}(ru\sigma)^2} \Phi\left(\frac{z}{\sigma} + ru\sigma\right) \\ &\quad + e^{rvz + \frac{1}{2}(rv\sigma)^2} \bar{\Phi}\left(\frac{z}{\sigma} + rv\sigma\right). \end{aligned}$$

Thus,  $\psi(z) = f_1(\alpha, z)f_2(z)$ . Because  $df_1(\alpha, z)/dz = -\alpha f_1(\alpha, z)$ ,

$$\psi'(z) = (-\alpha f_2(z) + f_2'(z))f_1(\alpha, z). \quad (7)$$

Because  $f_1(\alpha, z) > 0$  for all  $z$ , the sign of  $\psi'(z)$  is the same as the sign of  $-\alpha f_2(z) + f_2'(z)$ . Note that

$$\begin{aligned} f_2'(z) &= r u e^{ruz + \frac{1}{2}(ru\sigma)^2} \Phi\left(\frac{z}{\sigma} + ru\sigma\right) \\ &\quad + e^{ruz + \frac{1}{2}(ru\sigma)^2} \phi\left(\frac{z}{\sigma} + ru\sigma\right) \frac{1}{\sigma} \\ &\quad + r v e^{rvz + \frac{1}{2}(rv\sigma)^2} \bar{\Phi}\left(\frac{z}{\sigma} + rv\sigma\right) \\ &\quad - e^{rvz + \frac{1}{2}(rv\sigma)^2} \phi\left(\frac{z}{\sigma} + rv\sigma\right) \frac{1}{\sigma} \\ &= r u e^{ruz + \frac{1}{2}(ru\sigma)^2} \Phi\left(\frac{z}{\sigma} + ru\sigma\right) \\ &\quad + r v e^{rvz + \frac{1}{2}(rv\sigma)^2} \bar{\Phi}\left(\frac{z}{\sigma} + rv\sigma\right). \end{aligned}$$

Therefore,

$$\begin{aligned} -\alpha f_2(z) + f_2'(z) &= r(u - \alpha) e^{ruz + \frac{1}{2}(ru\sigma)^2} \Phi\left(\frac{z}{\sigma} + ru\sigma\right) \\ &\quad - r(\alpha - v) e^{rvz + \frac{1}{2}(rv\sigma)^2} \bar{\Phi}\left(\frac{z}{\sigma} + rv\sigma\right) \\ &= r e^{rvz} \left[ (u - \alpha) e^{r(u-v)z + \frac{1}{2}(ru\sigma)^2} \Phi\left(\frac{z}{\sigma} + ru\sigma\right) \right. \\ &\quad \left. - (\alpha - v) e^{\frac{1}{2}(rv\sigma)^2} \bar{\Phi}\left(\frac{z}{\sigma} + rv\sigma\right) \right]. \quad (8) \end{aligned}$$

Therefore, the sign of  $-\alpha f_2(z) + f_2'(z)$  is the same as the sign of the term inside the square brackets,

which is equal to an increasing function of  $z$  minus a decreasing function of  $z$ . Moreover, the term inside the square brackets is negative as  $z \rightarrow -\infty$  and positive as  $z \rightarrow +\infty$ . Therefore, the sign of  $-\alpha f_2(z) + f_2'(z)$  changes from negative to positive exactly once, implying that  $\psi(\cdot)$  is quasi-convex.

To prove that  $\psi(z) \geq 1$  for all  $z$ , note that the type- $\theta$  agent's expected utility satisfies the following inequalities:

$$\begin{aligned} E[-e^{-r(s(X|F)-V(a))}] &\leq -e^{-r(E[s(X|F)]-V(a))} \leq -e^{-r(E[\bar{s}(X)]-V(a))} \\ &= -e^{-r(\alpha(\theta+a)+\beta-V(a))}, \quad (9) \end{aligned}$$

where  $\bar{s}(x) = \alpha x + \beta$  because the utility function is concave (for the first inequality) and  $s(x|F) \leq \bar{s}(x)$  for all  $x$  and  $F$  (for the second inequality). As we saw earlier,  $E[-e^{-r(s(X|F)-V(a))}] = -e^{r(V(a)-\alpha a)} e^{-r\alpha\theta-r\beta} \psi(z)$ . Combining this inequality with (9), we have  $\psi(z) \geq 1$  for all  $z$ .  $\square$

**THEOREM 2.**  $z^*$  is increasing in  $\alpha$ , but decreasing in  $u$  and  $v$ .

**PROOF.** From (7) and the quasi convexity of  $\psi(\cdot)$ , we know  $z^*$  is the solution to

$$-\alpha f_2(z) + f_2'(z) = 0.$$

As  $\alpha$  increases to  $\alpha'$ ,  $-\alpha' f_2(z^*) + f_2'(z^*) < 0$ . To make this inequality an equality, the quasi convexity of  $\psi(\cdot)$  implies that  $z^*$  must be increased.

To show that  $z^*$  is decreasing in  $u$ , take the derivative of  $-\alpha f_2(z) + f_2'(z)$  with respect to  $u$  (using the expression in (8)):

$$r e^{ruz + \frac{1}{2}(ru\sigma)^2} \Phi\left(\frac{z}{\sigma} + ru\sigma\right) \quad (10)$$

$$+ r(u - \alpha)(rz + (r\sigma)^2 u) e^{ruz + \frac{1}{2}(ru\sigma)^2} \Phi\left(\frac{z}{\sigma} + ru\sigma\right) \quad (11)$$

$$+ r(u - \alpha) e^{ruz + \frac{1}{2}(ru\sigma)^2} \phi\left(\frac{z}{\sigma} + ru\sigma\right) (r\sigma). \quad (12)$$

The term (10) is clearly positive. Now consider the sign of the sum of (11) and (12) (and ignore the common term  $r(u - \alpha) e^{ruz + \frac{1}{2}(ru\sigma)^2}$  because it is positive). Note that

$$\begin{aligned} (rz + (r\sigma)^2 u) \Phi\left(\frac{z}{\sigma} + ru\sigma\right) + \phi\left(\frac{z}{\sigma} + ru\sigma\right) (r\sigma) \\ = (r\sigma) \left( \left( \frac{z}{\sigma} + ru\sigma \right) \Phi\left(\frac{z}{\sigma} + ru\sigma\right) + \phi\left(\frac{z}{\sigma} + ru\sigma\right) \right) \\ > 0, \end{aligned}$$

where the last inequality follows because  $w\Phi(w) + \phi(w) > 0$  for any  $w$ . (To see the latter, just show that the function of  $w$  is increasing and tends to 0 as

$w \rightarrow -\infty$ .) Therefore, as  $u$  increases,  $-r\alpha f_2(z) + f_2'(z)$  increases, implying that  $z^*$  is decreasing.

The fact that  $z^*$  is decreasing in  $v$  can be shown similarly. The derivative of  $-r\alpha f_2(z) + f_2'(z)$  with respect to  $v$  (using the expression in (8)) is

$$\begin{aligned} & r e^{rvz + \frac{1}{2}(rv\sigma)^2} \bar{\Phi}\left(\frac{z}{\sigma} + rv\sigma\right) \\ & - r(\alpha - v)(rz + (r\sigma)^2v) e^{rvz + \frac{1}{2}(rv\sigma)^2} \bar{\Phi}\left(\frac{z}{\sigma} + rv\sigma\right) \\ & + r(\alpha - v) e^{rvz + \frac{1}{2}(rv\sigma)^2} \phi\left(\frac{z}{\sigma} + rv\sigma\right) (r\sigma), \end{aligned}$$

which can be shown to be positive. (Here one uses the fact that  $\phi(w) - w\bar{\Phi}(w) > 0$  for any  $w$ . To see this, simply note that the function of  $w$  is decreasing and tends to zero as  $w \rightarrow +\infty$ .)  $\square$

It is interesting that the agent's forecast decision is analogous to the order quantity decision in a newsvendor problem. Suppose a retailer can buy a product for free and sell it for a per-unit price of  $\alpha$ . The retailer must determine a purchase quantity  $Q$  before the demand  $X$  is realized. The retailer's incentives for matching supply with demand come from the overage and underage costs: If supply exceeds demand, the retailer incurs a disposal cost of  $(u - \alpha)$  per unit of excess inventory; otherwise, if demand exceeds supply, the retailer must satisfy the excess demand by making an emergency purchase at a per-unit cost of  $(\alpha - v)$ . The retailer's profit is  $\alpha X - (u - \alpha) \cdot (Q - X)^+ - (\alpha - v)(X - Q)^+$ , which differs from  $s(X | Q)$  only by a constant  $\beta$ . Therefore, our agent's optimal forecast is the same as the retailer's optimal order quantity as long as the retailer is also risk averse with the same utility function. Risk-averse newsvendor problems have been studied by, e.g., Eeckhoudt et al. (1995) and Agrawal and Seshadri (2000). Sensitivity results similar to those in Theorem 2 have been reported in Eeckhoudt et al. (1995). Note that the agent's problem is only part of a larger problem, which is the principal's choice of contract parameters. Continuing with the above newsvendor analogy, we note that the values of  $\alpha$ ,  $u$ , and  $v$  determine the retailer's price/cost parameters. So, one can think of the principal's problem as one facing a manufacturer who can control the retailer's price/cost parameters. This kind of problem has often been addressed in the context of channel/supply chain coordination; see, e.g., Pasternack (1985), with risk neutrality at both levels being a common assumption. We now turn to the principal's problem.

## 5.2. The Principal's Problem

The principal's expected profit consists of three components: expected revenue  $E[X]$ , expected wage  $E[s(X | F)]$ , and the expected production/inventory

costs. Recall from the previous subsection that the agent's optimal effort is  $a^* = \alpha$  regardless of his type. Thus,  $X = \theta + \alpha + \epsilon$ . Because  $E[\epsilon] = 0$ ,  $E[X] = \bar{\theta} + \alpha$ . Also from the previous subsection, the agent's optimal forecast depends on his type: A low type leads to a forecast lower than the high type's. Therefore, the agent's forecast unambiguously signals the market condition. Under this separating case, the minimum expected production/inventory costs are  $G_1^*$  (see §2). It remains only to derive the expected wage paid to the agent.

Given that the agent is type  $\theta$ , forecasts  $F$ , and exerts effort  $a$ , the expected wage is

$$\begin{aligned} & E[s(X | F) | \theta, F, a] \\ & = E[s(\theta + a + \epsilon | F)] \\ & = \int_{-\infty}^F (s(F | F) - u(F - x)) \frac{1}{\sigma} \phi\left(\frac{x - \theta - a}{\sigma}\right) dx \\ & \quad + \int_F^{\infty} (s(F | F) + v(x - F)) \frac{1}{\sigma} \phi\left(\frac{x - \theta - a}{\sigma}\right) dx \\ & = \alpha F + \beta - u \int_{-\infty}^{z/\sigma} (z - \sigma y) \phi(y) dy \\ & \quad + v \int_{z/\sigma}^{\infty} (\sigma y - z) \phi(y) dy, \end{aligned}$$

where  $z = F - \theta - a$  and the last equality follows after a change of variable  $x = \sigma y + \theta + a$ . Note that

$$\int_{-\infty}^{z/\sigma} (z - \sigma y) \phi(y) dy = \sigma \left( \frac{z}{\sigma} \Phi\left(\frac{z}{\sigma}\right) + \phi\left(\frac{z}{\sigma}\right) \right)$$

and

$$\int_{z/\sigma}^{\infty} (\sigma y - z) \phi(y) dy = -\sigma \left( \frac{z}{\sigma} \bar{\Phi}\left(\frac{z}{\sigma}\right) - \phi\left(\frac{z}{\sigma}\right) \right).$$

Because  $F_H = \theta_H + \alpha + z^*$  and  $F_L = \theta_L + \alpha + z^*$ ,  $E[F] = \bar{\theta} + \alpha + z^*$ . The ex ante expected wage is

$$\begin{aligned} & \alpha(\bar{\theta} + \alpha + z^*) + \beta - u\sigma \left( \frac{z^*}{\sigma} \Phi\left(\frac{z^*}{\sigma}\right) + \phi\left(\frac{z^*}{\sigma}\right) \right) \\ & \quad - v\sigma \left( \frac{z^*}{\sigma} \bar{\Phi}\left(\frac{z^*}{\sigma}\right) - \phi\left(\frac{z^*}{\sigma}\right) \right) \\ & = \alpha(\bar{\theta} + \alpha) + (\alpha - v)z^* + \beta \\ & \quad - (u - v)\sigma \left( \frac{z^*}{\sigma} \Phi\left(\frac{z^*}{\sigma}\right) + \phi\left(\frac{z^*}{\sigma}\right) \right). \end{aligned}$$

There are two participation constraints, one for each agent type:

$$-e^{r(V(a^*) - \alpha a^*)} e^{-r(\alpha \theta_H + \beta)} \psi(z^*) \geq -U_0$$

and

$$-e^{r(V(a^*) - \alpha a^*)} e^{-r(\alpha \theta_L + \beta)} \psi(z^*) \geq -U_0,$$

where the agent's maximum expected utility for each type comes from the agent's objective function in (6),

with the decision variables  $a$  and  $z$  replaced with their optimal values. It is clear that the constraint for the high type is redundant. Recall that  $\beta$ , the fixed salary, does not impact the values of  $a^*$  and  $z^*$ . Therefore, the optimal  $\beta$  (from the firm's standpoint) binds the participation constraint for the low type. Because  $a^* = \alpha$  and  $V(a) = a^2/2$ , we have the optimal  $\beta$  (as a function of the other contract parameters):

$$\beta = -\frac{\alpha^2}{2} - \alpha\theta_L - \frac{\ln U_0}{r} + \frac{\ln \psi(z^*)}{r}.$$

Therefore, the minimum wage given  $\alpha$ ,  $u$ , and  $v$  is

$$\begin{aligned} & \alpha(\bar{\theta} + \alpha) - \frac{\alpha^2}{2} - \alpha\theta_L - \frac{\ln U_0}{r} + \frac{\ln \psi(z^*)}{r} + (\alpha - v)z^* \\ & - (u - v)\sigma \left( \frac{z^*}{\sigma} \Phi \left( \frac{z^*}{\sigma} \right) + \phi \left( \frac{z^*}{\sigma} \right) \right). \end{aligned} \quad (13)$$

Combining the above revenue and cost terms, we have the firm's expected profit:

$$\begin{aligned} & (\bar{\theta} + \alpha) - G_1^* - \left\{ \alpha(\bar{\theta} + \alpha) - \frac{\alpha^2}{2} - \alpha\theta_L - \frac{\ln U_0}{r} \right. \\ & \quad \left. + \frac{\ln \psi(z^*)}{r} + (\alpha - v)z^* \right. \\ & \quad \left. - (u - v)\sigma \left( \frac{z^*}{\sigma} \Phi \left( \frac{z^*}{\sigma} \right) + \phi \left( \frac{z^*}{\sigma} \right) \right) \right\}. \end{aligned} \quad (14)$$

The firm seeks the values of  $\alpha$ ,  $u$ , and  $v$  that maximize the above expression.

### 5.3. Bounds

This subsection presents bounds that are helpful in determining a numerical solution to the principal's contract-design problem. We first show that  $z^*$  is within easily computable bounds. These bounds, together with the quasi convexity of  $\psi(\cdot)$ , can be used for a quick calculation of  $z^*$ . We then show that the principal's objective function is bounded from above by a concave quadratic function of  $\alpha$ . This leads to bounds on the optimal value of  $\alpha$ . This subsection can be safely skipped in the first reading.

We first show that  $\psi(\cdot)$  can be bounded below by a simple function. Define  $s_u(x | F) = \alpha F + \beta - u(F - x)$  and  $s_v(x | F) = \alpha F + \beta + v(x - F)$  for all  $x$ . It is clear that  $s(x | F) = \min\{s_u(x | F), s_v(x | F)\}$ . Therefore,

$$\begin{aligned} E[e^{-rs(X|F)}] &= E[e^{-r \min\{s_u(X|F), s_v(X|F)\}}] \\ &= E[e^{\max\{-rs_u(X|F), -rs_v(X|F)\}}] \\ &= E[\max\{e^{-rs_u(X|F)}, e^{-rs_v(X|F)}\}] \\ &\geq \max\{E[e^{-rs_u(X|F)}], E[e^{-rs_v(X|F)}]\}. \end{aligned}$$

Suppose the agent is type  $\theta$ , forecasts  $F$ , and exerts effort  $a$ . Let  $z = F - \theta - a$ . Because  $X = \theta + a + \epsilon$ ,  $s_u(X | F) \sim N(\alpha(z + \theta + a) + \beta - uz, (u\sigma)^2)$ . Therefore,

the agent's certainty equivalent of the random income  $s_u(X | F)$  is  $\alpha(z + \theta + a) + \beta - uz - \frac{1}{2}r(u\sigma)^2$ . Similarly, the certainty equivalent of  $s_v(X | F)$  is  $\alpha(z + \theta + a) + \beta - vz - \frac{1}{2}r(v\sigma)^2$ . Consequently,

$$\begin{aligned} & E[e^{-rs(X|F)}] \\ & \geq \max\{e^{-r(\alpha(z+\theta+a)+\beta-uz-\frac{1}{2}r(u\sigma)^2)}, \\ & \quad e^{-r(\alpha(z+\theta+a)+\beta-vz-\frac{1}{2}r(v\sigma)^2)}\} \\ & = e^{-r(\alpha(\theta+a)+\beta)} \max\{e^{r(u-\alpha)z+\frac{1}{2}(ru\sigma)^2}, e^{r(v-\alpha)z+\frac{1}{2}(rv\sigma)^2}\}. \end{aligned}$$

On the other hand, replace  $F$  in (5) with  $z + \theta + a$ , and we have

$$E[e^{-rs(X|F)}] = e^{-r(\alpha(\theta+a)+\beta)}\psi(z).$$

Consequently,

$$\begin{aligned} & \psi(z) \geq \psi_{lb}(z) \\ & \stackrel{\text{def}}{=} \max\{e^{r(u-\alpha)z+\frac{1}{2}(ru\sigma)^2}, e^{r(v-\alpha)z+\frac{1}{2}(rv\sigma)^2}\}; \end{aligned} \quad (15)$$

i.e.,  $\psi(\cdot)$  is bounded below by the maximum of two simple convex functions of  $z$ , one increasing in  $z$  and the other decreasing.

The immediate benefits of the above lower-bound function are easily computable bounds around  $z^*$ . Let  $y_0$  be the value of  $\psi(z)$  for any given  $z$ . For example,  $y_0 = \psi(0) = e^{\frac{1}{2}(ru\sigma)^2}\Phi(ru\sigma) + e^{\frac{1}{2}(rv\sigma)^2}\bar{\Phi}(rv\sigma)$ . Let  $\bar{z} = \max\{z: \psi_{lb}(z) \leq y_0\}$  and  $\underline{z} = \min\{z: \psi_{lb}(z) \leq y_0\}$ . From (15),  $\underline{z} \leq z^* \leq \bar{z}$ . To obtain  $\bar{z}$  and  $\underline{z}$ , simply solve

$$e^{r(u-\alpha)\bar{z}+\frac{1}{2}(ru\sigma)^2} = y_0 \quad \text{and} \quad e^{r(v-\alpha)\underline{z}+\frac{1}{2}(rv\sigma)^2} = y_0$$

to arrive at

$$\bar{z} = \frac{\ln y_0 - \frac{1}{2}(ru\sigma)^2}{r(u-\alpha)} \quad \text{and} \quad \underline{z} = \frac{\ln y_0 - \frac{1}{2}(rv\sigma)^2}{r(v-\alpha)}.$$

The minimum point of  $\psi(z)$ ,  $z^*$ , can then be obtained by, e.g., a bisection search over  $z \in [\underline{z}, \bar{z}]$ .

Next, we derive an upper bound on the firm's expected profit. First, fix  $\alpha$  and consider the firm's expected profit as a function of  $u$  and  $v$ . (Recall that  $\beta$  is always set at its minimum value, which binds the participation constraint for the low-type agent. In other words,  $\beta$  has been replaced by a function of the other contract parameters.) Note that the firm's expected revenue and expected operations costs are independent of  $u$  and  $v$ , and only the agent's expected wage depends on  $u$  and  $v$ . Consider the expected wage in (13) after removing all the terms that are not affected by  $u$  and  $v$ :

$$\frac{\ln \psi(z^*)}{r} + (\alpha - v)z^* - (u - v)\sigma \left( \frac{z^*}{\sigma} \Phi \left( \frac{z^*}{\sigma} \right) + \phi \left( \frac{z^*}{\sigma} \right) \right). \quad (16)$$

(Note that  $z^*$  depends on  $u$  and  $v$ .) To derive a lower bound on the above expression (and thus eventually an upper bound on the firm's objective function), we replace  $\psi(\cdot)$  with  $\psi_{lb}(\cdot)$  and allow the firm to choose the value of  $z$ . That is, the solution to the following problem provides a lower bound on (16):

$$\min_z \frac{\ln \psi_{lb}(z)}{r} + (\alpha - v)z - (u - v)\sigma \left( \frac{z}{\sigma} \Phi \left( \frac{z}{\sigma} \right) + \phi \left( \frac{z}{\sigma} \right) \right). \quad (17)$$

To solve the problem in (17), recall that  $\psi_{lb}(\cdot)$  is the maximum of two convex functions, one increasing and the other decreasing. The two convex functions intersect at  $z_0$ , which is the solution to

$$r(u - \alpha)z + \frac{1}{2}(ru\sigma)^2 = r(v - \alpha)z + \frac{1}{2}(rv\sigma)^2.$$

Thus,  $z_0 = -\frac{1}{2}r\sigma^2(u + v)$ . For  $z \geq z_0$ , the objective function in (17) can be written as

$$\begin{aligned} & (u - \alpha)z + \frac{1}{2}r(u\sigma)^2 + (\alpha - v)z \\ & - (u - v)\sigma \left( \frac{z}{\sigma} \Phi \left( \frac{z}{\sigma} \right) + \phi \left( \frac{z}{\sigma} \right) \right) \\ & = (u - v)z - (u - v)\sigma \left( \frac{z}{\sigma} \Phi \left( \frac{z}{\sigma} \right) + \phi \left( \frac{z}{\sigma} \right) \right) + \frac{1}{2}r(u\sigma)^2 \end{aligned}$$

because  $\psi_{lb}(z) = e^{r(u-\alpha)z + \frac{1}{2}(ru\sigma)^2}$  for  $z \geq z_0$ . It is easy to verify that the above function is increasing in  $z$ . Similarly, the objective function is decreasing for  $z < z_0$ . Therefore, the solution to (17) is  $z = z_0$ , with the minimum objective function value

$$\begin{aligned} & (u - v)z_0 - (u - v)\sigma \left( \frac{z_0}{\sigma} \Phi \left( \frac{z_0}{\sigma} \right) + \phi \left( \frac{z_0}{\sigma} \right) \right) + \frac{1}{2}r(u\sigma)^2 \\ & = \frac{1}{2}r(\sigma v)^2 - (u - v)\sigma \\ & \cdot \left( \frac{-\frac{1}{2}r\sigma^2(u + v)}{\sigma} \Phi \left( \frac{-\frac{1}{2}r\sigma^2(u + v)}{\sigma} \right) \right. \\ & \quad \left. + \phi \left( \frac{-\frac{1}{2}r\sigma^2(u + v)}{\sigma} \right) \right). \quad (18) \end{aligned}$$

Now consider the problem of minimizing (18) over  $u$  and  $v$ . If we keep  $u + v$  constant and decrease  $v$  (thus increase  $u$ ), the above expression decreases because  $w\Phi(w) + \phi(w) > 0$  for all  $w$ . Setting  $v = 0$ , we have a new lower bound:

$$-u\sigma \left( -\frac{1}{2}r\sigma u \Phi \left( -\frac{1}{2}r\sigma u \right) + \phi \left( -\frac{1}{2}r\sigma u \right) \right).$$

Letting  $w = -\frac{1}{2}r\sigma u$ , the above lower bound can be written as

$$\frac{2}{r}w(w\Phi(w) + \phi(w)). \quad (19)$$

**LEMMA 1.** Define  $f(w) = w(w\Phi(w) + \phi(w))$  for all  $w$ . Then,  $f(\cdot)$  is quasi-convex. Moreover,  $f(w)$  is minimized at  $w \approx -0.6$  with  $f(-0.6) \approx -0.1$ .

**PROOF.** Note that

$$f'(w) = 2w\Phi(w) + \phi(w) = 2\phi(w) \left( \frac{w\Phi(w)}{\phi(w)} + \frac{1}{2} \right).$$

To examine the sign of  $f'(w)$ , one only needs to look at the sign of the term inside the brackets. Note that

$$\left( \frac{w\Phi(w)}{\phi(w)} \right)' = \frac{\Phi(w) + w\phi(w) + w^2\Phi(w)}{\phi(w)} > 0$$

because  $\Phi(w) + w\phi(w) + w^2\Phi(w) > 0$  for all  $w$ . (To see this, simply note that  $(\Phi(w) + w\phi(w) + w^2\Phi(w))' = 2\phi(w) + 2w\Phi(w) > 0$ , a fact noted earlier, and that  $\lim_{w \rightarrow -\infty} \Phi(w) + w\phi(w) + w^2\Phi(w) = 0$ .) Also, one can easily find a  $w$  value with  $f'(w) < 0$  and a  $w$  value with  $f'(w) > 0$  (such as  $w = 0$ ). Therefore  $f(\cdot)$  changes from negative to positive exactly once, implying that  $f(\cdot)$  is quasi-convex. Numerical optimization shows that  $f(w)$  is minimized at  $w \approx -0.6$  with  $f(-0.6) \approx -0.1$ .  $\square$

The above lemma implies that a lower bound on (19) is  $-0.2/r$ . Combining this lower bound on (16) with those terms in the expected wage expression (13) that were removed because they are independent of  $u$  and  $v$ , we have a lower bound on the expected wage (13):

$$\alpha(\bar{\theta} + \alpha) - \frac{\alpha^2}{2} - \alpha\theta_L - \frac{\ln U_0}{r} - \frac{0.2}{r}.$$

Combining the above expression with the firm's expected revenue and expected production/inventory costs, we have an upper bound on the firm's profit function:

$$\begin{aligned} \bar{\pi}(\alpha) & \stackrel{\text{def}}{=} (\bar{\theta} + \alpha) - G_1^* \\ & - \left\{ \alpha(\bar{\theta} + \alpha) - \frac{\alpha^2}{2} - \alpha\theta_L - \frac{\ln U_0}{r} - \frac{0.2}{r} \right\} \\ & = (1 - \alpha)(\bar{\theta} + \alpha) - G_1^* + \frac{\alpha^2}{2} + \alpha\theta_L + \frac{\ln U_0}{r} + \frac{0.2}{r}. \end{aligned}$$

This upper bound is a concave, quadratic function of  $\alpha$ . Let  $\pi_0$  be a feasible value of the firm's expected profit. Define  $\underline{\alpha} = \min\{\alpha: \bar{\pi}(\alpha) \geq \pi_0\}$  and  $\bar{\alpha} = \max\{\alpha: \bar{\pi}(\alpha) \geq \pi_0\}$ . The following result is immediate.

**THEOREM 3.**  $\underline{\alpha} \leq \alpha^* \leq \bar{\alpha}$ , where  $\alpha^*$  is the optimal value of  $\alpha$ .

#### 5.4. Optimization

The above results suggest the following algorithm for determining the values of the contract parameters  $\alpha$ ,  $\beta$ ,  $u$ , and  $v$  that maximize the firm's expected profit. As noted earlier, the optimal value of  $\beta$  is uniquely determined by the other three contract parameters through the binding participation constraint of the low-type agent. It remains only to determine the optimal values of  $\alpha$ ,  $u$ , and  $v$ . To this end, first compute

the bounds  $\underline{\alpha}$  and  $\bar{\alpha}$  on the optimal  $\alpha$ . The algorithm has two main loops. The outer loop iterates through a finite number of values of  $\alpha$  in the interval  $[\underline{\alpha}, \bar{\alpha}]$ , determined by a given step size. For each value of  $\alpha$ , the inner loop performs a grid search over  $(u, v)$ . For each given pair  $(u, v)$  inside the inner loop, first determine  $z^*$  and then use this value to determine the optimal corresponding value of  $\beta$ , as well as the firm's expected profit. The optimal values of  $\alpha, u,$  and  $v$  are the ones that return the maximum value of (14), the firm's expected profit. Note that  $G_1^*$  can be computed outside of the two loops, because it does not affect the optimal values of the contract parameters.

### 6. Numerical Examples

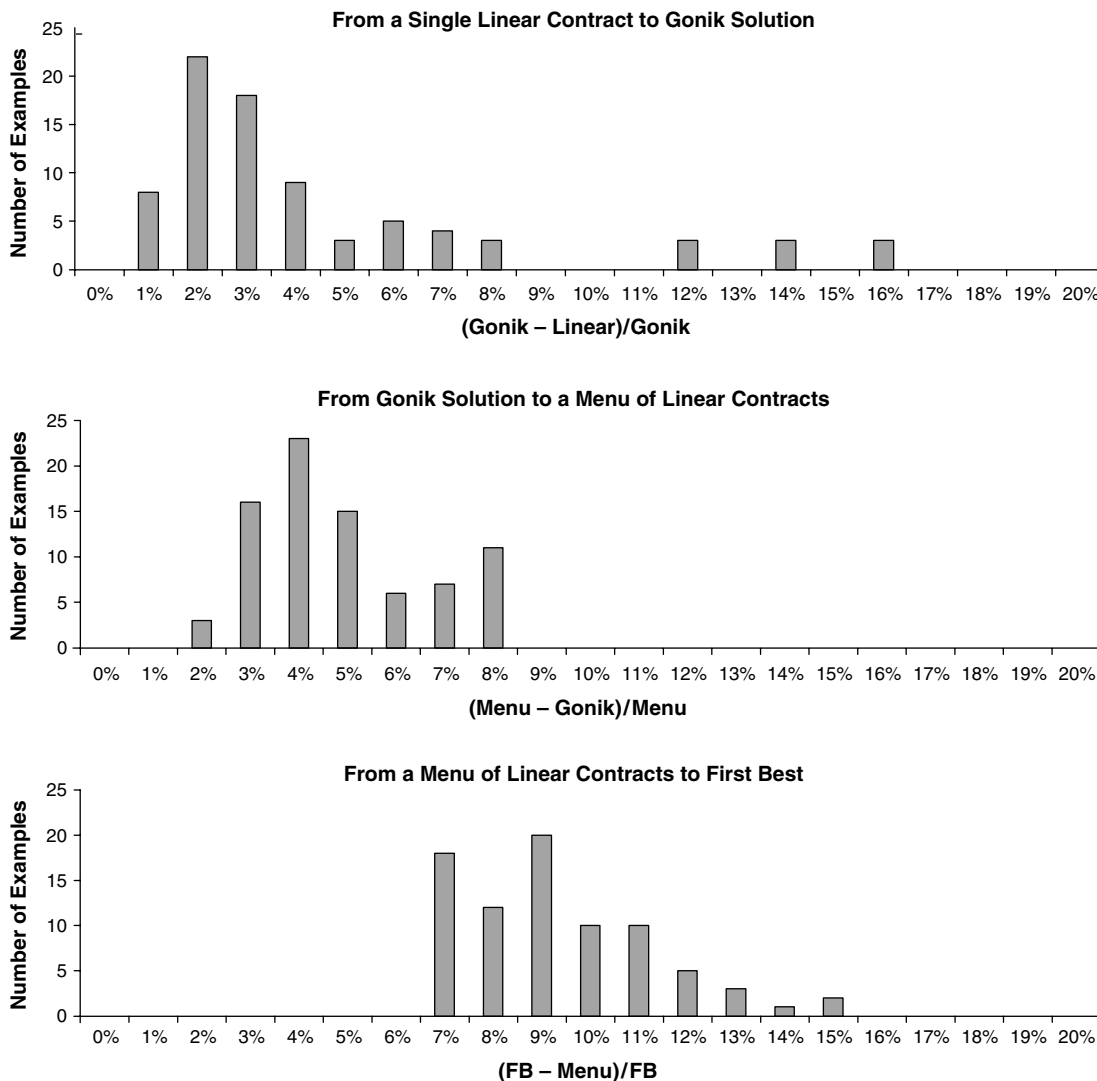
Here we numerically compare the Gonik solution with a menu of linear contracts, using as benchmarks the first-best solution and the pooling solution under a single linear contract. We fixed some parameters

and varied the others to arrive at 81 examples:  $\rho = 0.5$ ;  $U_0 = 1$ ;  $p = 0.1$ ;  $c' = 1$ ;  $(\theta_L, \theta_H) = (2, 3), (3, 4), (2, 5)$ ;  $\sigma = 0.2, 0.5, 0.8$ ;  $r = 0.5, 1, 2$ ; and  $c = 0.2, 0.5, 0.8$ . The three pairs of  $(\theta_L, \theta_H)$  are indexed by 1, 2, and 3 in Figure 3.

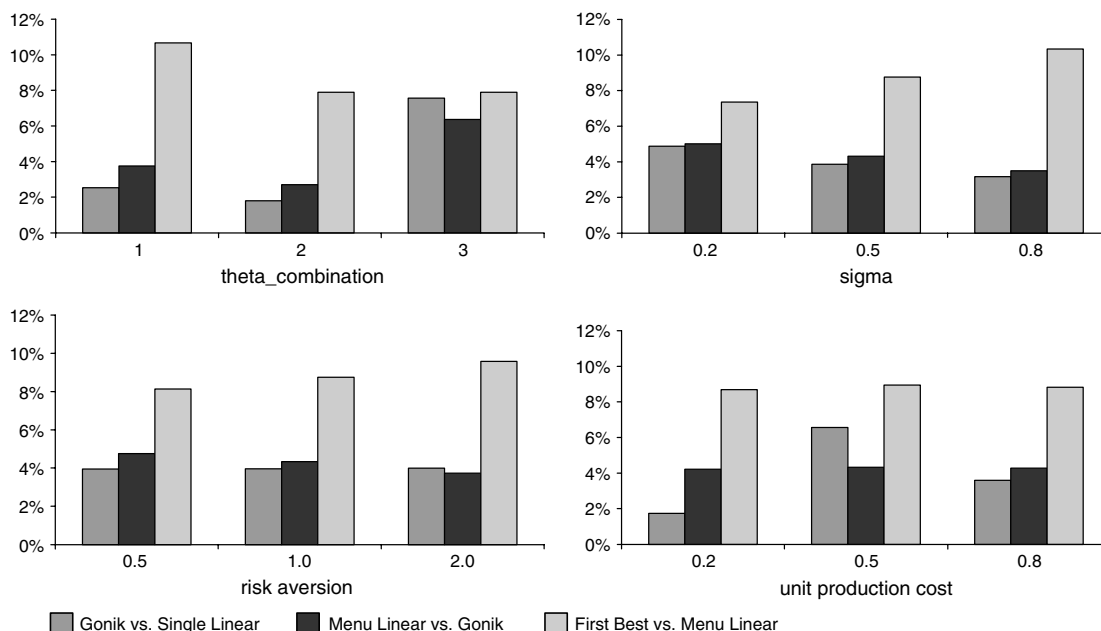
For each example, we computed the firm's maximum expected profits under the first-best solution, a single linear contract, a menu of linear contracts, and the Gonik scheme. The relative differences among these profits are tabulated and reported in Figure 2. Figure 3 then examines these differences more closely, seeking relationships, if any, between the quality of a solution and model parameters. The numerical results lead to the following observations.

First, the Gonik solution dominates the pooling solution, and the average percentage profit difference is 3.97%. This shows that the firm benefits from a menu contract that leads to the separation of the agent types. (Put differently, the market information

Figure 2 Comparisons of Firm Profits



**Figure 3 Profit Gaps and System Parameters**



is valuable.) The gap increases when there is greater variability in the market conditions (reflected in a larger  $\theta_H - \theta_L$ ), the demand variance decreases, or the unit production cost is close to the middle between  $p$  and  $c'$ . The gap is insensitive to the agent's degree of risk aversion.

Second, a menu of linear contracts dominates the Gonik solution, with an average percentage profit difference of 4.27%. The gap between these two solutions increases when (a) there is greater variability in the market conditions, (b) the variance of the random noise term decreases, and (c) the agent becomes less risk averse. Recall that under Gonik's scheme, both types of agent exert the same level of effort, whereas a menu of linear contracts calls for a higher level of effort for the high type. It is perhaps this flexibility that makes a menu of linear contracts a better solution. As  $\theta_H - \theta_L$  increases, the abovementioned flexibility becomes more important, explaining (a).

Third, the gap between a menu of linear contracts and the first-best solution is significant, with an average percentage difference of 8.82%. The gap increases when (i) the market contracts (reflected in a smaller  $\theta$  value), (ii) the variance of the random noise term increases, and (iii) the agent becomes more risk averse. Note that (ii) and (iii) are consistent with the basic moral hazard model: Larger  $\sigma$  means more noise in the measurement of the agent's effort, and more risk aversion means the provision of incentives becomes more costly for the principal; both of these point to a more serious moral hazard problem. The reason for (i) seems to be that as  $\bar{\theta}$  decreases, the firm's profit decreases under both solutions, amplifying their relative gap.

Finally, note that the unit production cost  $c$  (and for that matter, the salvage value  $p$  and the overtime cost  $c'$ ) does not affect the firm's optimization problem for choosing the optimal contract parameters under the first-best solution, a menu of linear contracts, and the Gonik scheme. Therefore, a change in  $c$  will cause the firm's profits under these solutions to change by the same amount. The sensitivity analysis presented in Figure 3 over the unit production cost shows that the percentage profit differences (among the three solutions) become higher for  $c = 0.5$ . The reason is that it is often true that if the values of  $c - p$  and  $c' - c$  become closer, while keeping their total constant, the minimum inventory cost increases (this is easy to verify in the EOQ model with backorders). Note that one can arbitrarily increase (or decrease) the percentage profit differences between two contracting approaches by increasing (or decreasing) the operating costs.

## 7. Concluding Remarks

Gonik (1978) said it well when he described the objectives all managers are shooting for in a complex sales environment:

First, they want to pay salesmen for their absolute sales volume. Second, they want to pay them for their effort, even if they are in tough areas where they will sell less. Third, they want good and fresh field information on market potential for planning and control purposes. (p. 118)

A clever method to achieve these goals, the Gonik scheme is widely quoted in both the marketing and

the economics/agency literature. The contribution of this paper is, then, (1) to devise a model that blends the above first and second challenges articulated by Gonik with a concrete benefit of good and fresh field information in production and inventory planning, (2) to provide a careful analysis of the Gonik scheme in the context of this model, and (3) to compare the Gonik scheme with a menu of linear contracts, a contemporary solution often suggested in the agency literature. Our analysis shows that the Gonik scheme is dominated by a menu of linear contracts.

As mentioned earlier, the Gonik scheme induces the same level of effort for both agent types, whereas a menu of linear contracts can cause the agent to exert different levels of effort under different market conditions. One can modify the Gonik scheme so as to induce different levels of effort from different agent types. Consider  $s(F | F) = f_0 + f_1F + f_2F^2$  for some constants  $f_0, f_1, f_2$ . (The original Gonik scheme has  $f_2 = 0$ .) Using this new form of  $s(F | F)$  in (4), we have a new menu of nonlinear contracts. Call this the new Gonik scheme. A full-blown analysis of this contract form is complex. However, if we assume that  $u = v = s'(F | F) = f_1 + 2f_2F$ , then the analysis is fairly straightforward. In this case, the firm offers a menu of linear contracts, and if the agent submits a forecast  $F$ , he has effectively chosen a linear contract  $s(x | F) = \alpha(F)x + \beta(F)$ , where  $\alpha(F) = f_1 + 2f_2F$  and  $\beta(F) = s(F | F) - \alpha(F)F = f_0 - f_2F^2$ . For this new Gonik scheme, it can be shown that the different agent types will indeed choose different linear contracts and exert different levels of effort. An immediate question is whether or not this new scheme will outperform the menu of linear contracts considered in §4. The answer is no. This is simply because the two linear contracts chosen by the two agent types under the new Gonik scheme can be offered directly to the agent at the outset (i.e., without going through the  $s(F | F)$  calculations). The question is how much of the gap between the original Gonik scheme and a menu of linear contracts (as considered in §4) can be closed by the new Gonik scheme—according to preliminary numerical analysis, not much. The reason appears to be that the new Gonik scheme offers an infinite number of linear contracts, and these contracts are not unrelated because they must be tangents of the same quadratic function. This gives the agent much flexibility at the expense of the principal. However, this disadvantage of the new Gonik scheme is likely to diminish as the number of possible market conditions increases. Of course, the Gonik idea may still prevail with more general functional forms, and/or with  $u$  and  $v$  chosen optimally; we will leave this for future research. (The analysis of the quadratic Gonik scheme can be obtained from the author upon request.)

Although our numerical comparison shows that the Gonik scheme is dominated by a menu of linear contracts, there are other aspects that need to be considered before a definitive conclusion can be drawn about these two solutions. This paper has assumed a single sales agent serving a single sales territory that has two possible market conditions (high or low). Reality is always much more complicated. As the sales environment gets more complicated, it seems that the Gonik scheme has the ability to keep the contract simple, and thus easily implementable, as compared with a menu of linear contracts. For example, with each additional sales territory, the number of contracting parameters increases by one under the Gonik scheme (the firm needs to provide a quota  $O_i$  for territory  $i$  and compensates agent  $i$  by the amount  $s(X_i/O_i | F_i/O_i)$ , where  $X_i$  is the actual sales in territory  $i$  and  $F_i$  is the forecast by agent  $i$ ), whereas under a menu of linear contracts, the number of contracting parameters (the slope and intercept of each linear function on the menu) increases by two times the number of different possible market conditions in the added territory. By this count, the Gonik scheme is simpler than a menu of linear contracts. This advantage of the Gonik scheme must be considered together with the firm's profit performance to say which of these solutions is truly better in a multiterritory sales environment.

### Acknowledgments

The author would like to acknowledge financial support from the Columbia Business School, the National Science Foundation, and the Stanford Global Supply Chain Management Forum. Thanks to Bob Gibbons of Massachusetts Institute of Technology, two anonymous referees, and the seminar participants at Columbia University, the INFORMS 2000 meeting in Salt Lake City, the Stanford Workshop on Incentives and Coordination in Operations Management, the University of British Columbia, and the University of California at Berkeley for helpful suggestions and comments. Aradhna Krishna of the University of Michigan first brought the Gonik article in *Harvard Business Review* to the author's attention. The first draft of this paper was written while the author was on sabbatical leave at Stanford University. Thanks to the Operations Information and Technology group of the Stanford Business School for their hospitality.

### References

- Agrawal, V., S. Seshadri. 2000. Impact of uncertainty and risk aversion on price and order quantity in the newsvendor problem. *Manufacturing Service Oper. Management* 2(4) 410–423.
- Basu, A., R. Lal, V. Srinivasan, R. Staelin. 1985. Salesforce-compensation plans: An agency theoretic perspective. *Marketing Sci.* 4(4) 267–291.
- Chen, F. 2000. Salesforce incentives and inventory management. *Manufacturing Service Oper. Management* 2(2) 186–202.
- Coughlan, A. 1993. Salesforce compensation: A review of MS/OR advances. J. Eliashberg, G. Lilien, eds. *Handbooks in Operations Research and Management Science: Marketing*, Vol. 5. North-Holland, Amsterdam, The Netherlands.

- Dearden, J., G. Lilien. 1990. On optimal salesforce compensation in the presence of production learning effects. *Internat. J. Res. Marketing* 7(2–3) 179–188.
- Eeckhoudt, L., C. Gollier, H. Schlesinger. 1995. The risk-averse (and prudent) newsboy. *Management Sci.* 41(5) 786–794.
- Gibbons, R. 1987. Piece-rate incentive schemes. *J. Labor Econom.* 5(4) 413–429.
- Gonik, J. 1978. Tie salesmen's bonuses to their forecasts. *Harvard Bus. Rev.* 56(May–June) 116–122.
- Grossman, S., O. Hart. 1983. An analysis of the principal-agent problem. *Econometrica* 51(1) 7–45.
- Harris, M., A. Raviv. 1978. Some results on incentive contracts with applications to education and employment, health insurance, and law enforcement. *Amer. Econom. Rev.* 68(March) 20–30.
- Harris, M., A. Raviv. 1979. Optimal incentive contracts with imperfect information. *J. Econom. Theory* 20(April) 231–259.
- Holmstrom, B. 1979. Moral hazard and observability. *Bell J. Econom.* 10(1) 74–91.
- Holmstrom, B. 1982. Moral hazard in teams. *Bell J. Econom.* 13(2) 324–340.
- Kreps, D. 1990. *A Course in Microeconomic Theory*. Princeton University Press, Princeton, NJ.
- Laffont, J.-J., J. Tirole. 1986. Using cost observation to regulate firms. *J. Political Econom.* 94(3) 614–641.
- Lal, R. 1986. Delegating pricing responsibility to the salesforce. *Marketing Sci.* 5(2) 159–168.
- Lal, R., R. Staelin. 1986. Salesforce-compensation plans in environments with asymmetric information. *Marketing Sci.* 5(3) 179–198.
- Mantrala, M. K., K. Raman. 1990. Analysis of a sales force incentive plan for accurate sales forecasting and performance. *Internat. J. Res. Marketing* 7 189–202.
- Pasternack, B. 1985. Optimal pricing and return policies for perishable commodities. *Marketing Sci.* 4(2) 166–176.
- Porteus, E., S. Whang. 1991. On manufacturing/marketing incentives. *Management Sci.* 37(9) 1166–1181.
- Raju, J., V. Srinivasan. 1996. Quota-based compensation plans for multiterritory heterogeneous salesforces. *Management Sci.* 42(10) 1454–1462.
- Rao, R. 1990. Compensating heterogeneous salesforces: Some explicit solutions. *Marketing Sci.* 9(4) 319–341.
- Shavell, S. 1979. Risk sharing and incentives in the principal and agent relationship. *Bell J. Econom.* 10(1) 55–73.