Technical Note: Optimal Salesforce Compensation with Supply–Demand Mismatch Costs

Binqing Xiao
School of Management and Engineering, Nanjing University, Nanjing, Jiangsu 210093, China, bengking@nju.edu.cn

Wenqiang Xiao*
Stern School of Business, New York University, New York, New York 10012, USA, wxiao@stern.nyu.edu

In this study, we characterize the optimal compensation scheme for a firm that sells a single product with a limited stocking quantity through a sales agent. Our focus is on understanding how the supply–demand mismatch costs affect the firm’s optimal compensation scheme. There are two main findings. First, under the deterministic demand response, the classical optimality result of the convex increasing compensation scheme breaks with the consideration of supply–demand mismatch costs. Instead, the optimal compensation is S-shaped under certain conditions. Second, under the stochastic demand response, the classical optimality result of the menu of linear compensation schemes fails to hold with the consideration of supply–demand mismatch costs. Instead, the optimal compensation schemes consist of a menu of linear compensation coupled with a penalty of the agent’s forecast error.

Key words: salesforce compensation; S-shaped; supply–demand mismatch costs

History: Received: November 2017; Accepted: July 2019 by Albert Ha, after three revisions.

1. Introduction

There are three commonly used compensation plans for salespersons in practice, namely the linear plan (i.e., compensation is proportional to the sales achieved), the convex plan (i.e., increasing marginal compensation with increasing sales) and the S-shaped plan (i.e., increasing marginal compensation for sales below a threshold and decreasing marginal compensation for sales above that threshold). (Please see Lal and Srinivasan 1993, O’Connell 1989, and Oyer 2000 for a more detailed discussion on the use of these plans in practice.) Research on sales compensation has identified conditions under which specific plans are optimal - for example, Holmstrom and Milgrom (1987) discuss the optimality of the linear plan while Basu et al. (1985) discuss the optimality of the convex plan. These papers exclusively look at the economic trade-off between increasing revenues through increased sales and increased cost of effort for generating the increased sales. In this study, we include a third economic factor—namely, operational costs, as represented by the cost of having a mismatch between available inventory and demand—and show that the other compensation plan used in practice, that is, the S-shaped plan, is optimal under certain conditions.

More specifically, we examine the issue of how a firm should design the compensation scheme for a salesperson when the stocking decision for a short life cycle product has to be made in the face of demand uncertainty. We present a stylized model of this situation using the principal agent framework with both hidden information and hidden action. Our model differs from the classical agency models in that we take the supply–demand mismatch costs into consideration. Our focus is on understanding how the supply–demand mismatch costs impact the structure of the firm’s optimal compensation scheme. There are two main findings. First, under the deterministic demand response, we show that without the mismatch costs, the firm should offer increasing marginal rewards for increasing sales. In contrast, with the mismatch costs, the convex increasing property may no longer hold. Instead, we show that the optimal compensation scheme is S-shaped under certain conditions. Second, under the stochastic demand response, we show that without the mismatch costs, the firm should offer a menu of linear compensation schemes (MLC) under which the commission rate increases in the agent’s private demand condition. In contrast, with the mismatch costs, the convexity result of MLC breaks. Instead, we show that a menu of linear compensation coupled with a penalty of the agent’s forecast error.
Designing an effective compensation scheme is an important issue in salesforce management. In the marketing literature, Basu et al. (1985) is the first paper to employ the agency model to characterize the optimal compensation scheme in a setting where the principal and the agent have symmetric information about sales uncertainties before contracting. They derive conditions under which the optimal compensation function is convex in the sales. Rao (1990) considers an asymmetric information setting where the agent possesses superior knowledge about sales uncertainties than the principal. His main findings are as follows: (1) When the agent has perfect knowledge about sales uncertainties, the optimal compensation is a convex function of sales; (2) When the agent has imperfect knowledge about sales uncertainties, a menu of linear compensation plans is optimal for the principal. See Coughlan (1993) for a survey of this stream of literature. Our work is distinct from this stream of literature in that the principal must decide the inventory level before the season and the mismatch between supply and demand is costly to the principal. Interestingly, we show that with the presence of supply–demand mismatch costs, the optimal compensation scheme is S-shaped (i.e., first convex and then concave) under deterministic demand and the MLC with forecast penalties is optimal under stochastic demand. This suggests that the consideration of operational costs can have a significant impact on the structure of the optimal salesforce compensation scheme.

Our work is related to the literature on the Marketing/Operations interface that studies the question of how production-inventory planning impacts the design of salesforce compensation. Chen (2005) compares a forecast-based compensation scheme with a menu of linear contracts in a model where the salesperson has private information about the market condition, and this private information is useful for the firm in production planning. Sohoni et al. (2011) compare the piece-wise convex compensation scheme and the quota-based compensation scheme with the emphasis on both the expected sales and the sales variance. Khanjari et al. (2014) study the performance of a menu of linear contracts in a supply chain with either a retailer- or manufacturer-employed sales agent. Our work differs from these papers in that while they restrict attention to specific contract forms (such as quota-based plans), we consider a broader contract space and derive the optimal compensation scheme. As a result, while their focus is on how the operational factors impact the comparison between several classes of compensation schemes, our focus is on how the operational factors impact the shape of optimal compensation scheme. Chu and Lai (2013) and Dai and Jerath (2013) show that the salesquota-based bonus contract is optimal under demand censorship and symmetric information. We complement their work by showing that under asymmetric information the optimal compensation scheme is S-shaped under deterministic demand response and MLC coupled with forecast penalties under stochastic demand response.

2. Model

Consider a firm (she) selling a single product with finite inventory level $q$ through a sales agent (he) over a single selling season. The market demand in the selling season, denoted by $x$, is determined by both the market condition $\theta$ and the agent’s sales effort $e$, in the following additive form: $x = \theta + e$. The market condition $\theta$ is a random variable with the distribution function $F(\cdot)$ and the density function $f(\cdot)$ defined over the support $[\underline{\theta}, \overline{\theta}]$. The additive demand response has been commonly adopted in the agency compensation literature (see e.g., Chen 2005, Laffont and Tirole 1986). It is worth mentioning that we have verified that all of our core results (i.e., the optimality results given in Propositions 3 and 4) qualitatively hold for the multiplicative demand response as well (i.e., $x = \theta e$).

While the distribution function $F(\cdot)$ is common information for the firm, only the agent can privately observe the actual value of $\theta$ before the season due to his proximity to customers. This causes the classical hidden information problem for the firm in compensating the agent. Further, the agent’s sales effort $e$ is unobservable to the firm, causing the classical hidden action problem. For notational simplicity, we assume that the agent’s cost of effort $e$ is $C(e) = e^2/2$, with the remark that our core results hold for any convex increasing effort cost function.

Although the combined hidden information and hidden action problems have been examined in the salesforce and agency compensation literature (see, e.g., Laffont and Tirole 1986, Rao 1990), our model differs from this stream of literature in that we take the cost of supply–demand mismatch into the firm’s consideration. Specifically, given the supply quantity $q$ and the demand quantity $x$, the firm’s profits excluding the cost of compensation to the agent, denoted by $\Pi(q,x)$, are equal to

$$\Pi(q,x) = px - \Gamma(x - q),$$

where $p$ is the marginal profit excluding the mismatch cost and $\Gamma(\cdot)$ is the mismatch cost function. In the special case of $\Gamma(\cdot) \equiv 0$, our model reduces to the classical agency model with both hidden information and hidden action (see, e.g., Rao 1990) and without inventory consideration.
Our general model of mismatch cost includes the classical single-period inventory model with backorder (see, e.g., Chen 2005) as a special case. To see this, consider the backorder model where the firm first produces quantity \( q \) at per unit cost \( c \), and then demand \( x \) is realized. Let \( r \) be the per unit selling price. If the supply exceeds the demand (i.e., \( q > x \)), then the excess supply \( q-x \) is salvaged at per unit price \( s \); whereas if the supply is not enough to satisfy the demand (i.e., \( q < x \)), then the excess demand is satisfied by emergency (or expedited) production or backordering at per unit cost \( c_e \). Given the supply quantity \( q \) and the demand quantity \( x \), the firm’s profits \( \Pi(q,x) \) can be written as \( \Pi(q,x) = rx - cq - c_e (x-q)^+ + s(q-x)^+ = px - \Gamma(x-q) \), where \( p = r-c \), \( \Gamma(z) = (c_e-c)z^+ + (c-s)(-z)^+ \), and \( z^+ = \max(z,0) \).

Our goal is to examine how the consideration of supply–demand mismatch cost impacts the firm’s optimal compensation scheme. It follows from the revelation principle that without loss of optimality, the firm can restrict to the direct mechanism, denoted by \( (x(\theta),t(\theta)) \), with \( x(\theta) \) being the sales target and \( t(\theta) \) being the compensation intended for the agent who observed \( \Theta = \theta \) for \( \theta \in [\underline{\theta}, \bar{\theta}] \).

The sequence of events is summarized as follows. First, the firm stocks \( q \) units of inventory to be sold over a single selling season due to a relatively long production/order lead time, and offers the agent a direct mechanism \( (x(\theta),t(\theta)) \), for \( \theta \in [\underline{\theta}, \bar{\theta}] \). Second, the agent observes the realized value of the market condition \( \Theta = \theta \) and then chooses his intended sales-payment pair \( (x(\theta),t(\theta)) \) provided that it is in the best interest of the agent to do so. Third, the agent exerts effort \( e \) to achieve the sales target \( x(\theta) \) and receives the compensation \( t(\theta) \) from the firm who collects the sales profits (excluding compensation) \( \Pi(q,x(\theta)) \).

Consistent with the agency literature that characterizes the optimal compensation plan without consideration of mismatch costs (see, e.g., Laffont and Tirole 1986, Rao 1990), we assume that both the firm and the agent are risk neutral, each maximizing their own expected profits. While we can allow the stocking quantity \( q \) to be a decision variable for the firm, our core results as to how the supply–mismatch costs impact the shape of the firm’s optimal compensation scheme holds for any fixed \( q \) and hence also for the firm’s optimal stocking quantity. Therefore, in the remainder of the study, we assume \( q \) is fixed for exposition simplicity.

We make the following mild assumptions on the distribution of the market condition and the mismatch cost function. Define \( H(\theta) = (1-F(\theta))/f(\theta) \) which is the reciprocal of the hazard rate function of \( \Theta \), and we assume \( H(\theta) \) is a decreasing function of \( \theta \), that is, \( \Theta \) has an increasing hazard rate. The increasing hazard rate assumption is common in agency models with asymmetric information and is satisfied by many common distributions such as uniform, exponential, normal, gamma, etc. We assume that \( \Gamma(\cdot) \) is convex to ensure the firm’s optimal compensation problem is well defined.

Because the agent knows the actual value of \( \Theta \), the demand response function \( x = \Theta + e \) implies that the agent faces deterministic demand in deciding how much effort to exert. We will analyze this deterministic demand case in section 3. The results in section 3 have useful implications on the more realistic setting in which the agent faces stochastic demand in making the effort decision. In section 4, we will relax the deterministic demand assumption and analyze a stochastic demand model.

### 3. The Optimal Compensation Scheme under Deterministic Demand

In this section, we first formulate the firm’s compensation problem under any given direct mechanism; we then present the optimal direct mechanism and show that the optimal direct mechanism can be equivalently implemented by offering the agent the optimal payment scheme \( T^*(x) \); we finally discuss how the presence of supply–mismatch cost impacts the shape of the optimal payment scheme \( T^*(x) \).

Take any direct mechanism \( (x(\theta),t(\theta)) \). Let \( \pi(\theta, \hat{\theta}) \) denote the profit obtained by the agent if he observes that the market condition is \( \theta \) (called the type-\( \theta \) agent) but chooses the sales-payment pair \( (x(\hat{\theta}), t(\hat{\theta})) \). To ensure that the final sales meets the chosen sales target \( x(\hat{\theta}) \), the type-\( \theta \) must exert sales effort \( e = x(\hat{\theta}) - \hat{\theta} \). Hence,

\[
\pi(\theta, \hat{\theta}) = t(\hat{\theta}) - (x(\hat{\theta}) - \hat{\theta})^2/2.
\]

To ensure truth telling from the agent of any type, we must have that for all \( \theta, \hat{\theta} \in [\underline{\theta}, \bar{\theta}] \),

\[
\pi(\theta, \hat{\theta}) \geq \pi(\hat{\theta}, \hat{\theta}),
\]

which indicates that it is in the best interest of the type-\( \theta \) agent to choose the sales-payment pair \( (x(\theta), t(\theta)) \) intended for his true type \( \theta \).

To ensure participation from the agent of any type, we must have that for all \( \theta \in [\underline{\theta}, \bar{\theta}] \),

\[
\pi(\theta, \theta) \geq 0,
\]

where we have normalized the agent’s reservation profits to be zero without loss of generality.

Under the direct mechanism \( (x(\theta),t(\theta)) \) ensuring truth telling from the agent, the firm’s expected
profits are $E_\Theta[\Pi(q, x(\Theta)) - t(\Theta)]$. Therefore, the firm’s optimal direct mechanism is the solution to the following problem, denoted by (P),
\[
\max_{(x(\cdot), t(\cdot))} E_\Theta[\Pi(q, x(\Theta)) - t(\Theta)]
\]
\[\text{s.t. IC and IR.}\]

**PROPOSITION 1.** The firm’s optimal direct mechanism $(x^*(\cdot), t^*(\cdot))$ that solves the problem (P) is:
\[
x^*(\cdot) = e^*(\cdot) + \theta, \\
t^*(\cdot) = [e^*(\cdot)]^2/2 + \int_0^{e^*(\cdot)} e^*(z)dz,
\]
where $e^*(\cdot)$ is the type-$\Theta$ agent’s optimal effort level satisfying
\[
e^*(\cdot) = \arg \max_{\epsilon} \{pe - \epsilon^2/2 - \Gamma(e + \theta - q) - H(\theta)e\}. \tag{1}
\]

Because the agent faces deterministic demand, the firm can set the sales target $x(\cdot)$ accordingly to induce the type-$\Theta$ agent to exert the firm’s most desired effort level. Further, the firm can set the compensation $t(\cdot)$ properly to ensure that it is in the best interest of the type-$\Theta$ agent to fulfill the sales target $x(\cdot)$. Therefore, the deterministic demand assumption simplifies the firm’s problem to determining the firm’s most desired effort level from the type-$\Theta$ agent for $\theta \in [\underline{\theta}, \bar{\theta}]$. Equation (1) in Proposition 1 provides the answer for the firm’s most desired effort levels. The first three terms inside the maximization correspond to sales profits, costs of sales effort, and costs of supply–demand mismatch, respectively, on which the agent’s effort has direct impacts. The last term $H(\theta)e$ captures the impact of the type-$\Theta$ agent’s effort on the profit surplus of those higher types to which the firm must yield due to their informational advantage. Consistent with the literature on adverse selection, the information rent term $H(\theta)e$, together with the assumption that $H(\theta)$ decreases in $\theta$, provides stronger downward pressure on the firm’s most desired effort level as $\theta$ decreases. We call it the market information effect.

Under the optimal direct mechanism $(x^*(\cdot), t^*(\cdot))$, the type-$\Theta$ agent exerts effort level $e^*(\cdot)$, achieves the sales target $x^*(\cdot)$, and receives the compensation $t^*(\cdot)$ from the firm. Because $x^*(\cdot)$ strictly increases in $\theta$ (analytically established in the proof of Proposition 1), we can define its inverse function $\theta^*(x)$ such that $x^*(\theta^*(x)) = x$. Let $T^*(x) \equiv t^*(\theta^*(x))$. The monotonic property of $x^*(\cdot)$ implies that the firm can offer to the agent the compensation scheme $T^*(x)$, where $T^*(x)$ is the payment made by the firm to the agent if the sales volume is $x$. Doing so induces the same responses from all types of agents as those under the optimal direct mechanism $(x^*(\cdot), t^*(\cdot))$, resulting in the same optimal expected profits for the firm.

In the remainder of this section, we study the structure of the firm’s optimal compensation scheme $T^*(x)$. By its definition, we have $dT^*(x)/dx = e^*(\theta^*(x))$. This, together with the property that $e^*(\cdot)$ strictly increases in $x$, implies that the sign of the second order derivative $d^2T^*(x)/dx^2$ is the same as that of $de^*(\theta)/d\theta$ at $\theta = \theta^*(x)$. Without considering the mismatch costs, that is, $\Gamma(\cdot) = 0$, the sole presence of the market information effect results in the monotonic increasing property of $e^*(\cdot)$, implying that the firm’s optimal compensation scheme is convex increasing in sales, that is, offering progressively higher commission rate for higher sales to induce higher effort level from the higher type agent. Consistent with the extant literature (see Laffont and Tirole 1986, Rao 1990), we formally establish the convex increasing property in the case of no mismatch costs in the following proposition.

**PROPOSITION 2.** Without the supply–demand mismatch cost, that is, $\Gamma(\cdot) = 0$, the firm’s most desired effort level $e^*(\cdot)$ increases in $\theta$, and the firm’s optimal compensation scheme $T^*(x)$ is convex increasing in the sales $x$.

Contrast emerges when the firm takes the supply–demand mismatch cost into consideration. In contrast to the stronger downward pressure on the effort level of the lower type agent, the presence of $\Gamma(\cdot)$ in Equation (1) provides stronger upward lift on the effort level of the lower type agent to bridge the gap between supply and demand. We call it the operational mismatch effect, which works in the opposite direction as that of the market information effect on the firm’s most desired effort level $e^*(\cdot)$. Consequently, with the presence of both the market information effect and the operational mismatch effect, the monotonic increasing property of $e^*(\cdot)$ may no longer hold, implying that the classical convex increasing property of the firm’s optimal compensation scheme may break. In the following proposition, we provide sufficient conditions under which the convexity result breaks and instead, the firm’s optimal compensation scheme is S-shaped.

**PROPOSITION 3.** With the consideration of supply–demand mismatch cost, if the third-order derivative $\Gamma''''(x) \geq 0$ and $H(\theta)$ is convex in $\theta$, then as $\theta$ increases, $e^*(\theta)$ first increases and then decreases, implying that the firm’s optimal compensation scheme $T^*(x)$ is S-shaped, that is, first convex increasing and then concave increasing in the sales $x$.

The convex decreasing property of $H(\theta)$ implies that the market information effect (i.e., downward...
pressure on the effort level) is the strongest at \( \theta = \hat{\theta} \)
and gradually weakens as \( \theta \) increases. Therefore, for
the low values of \( \theta \), the downward pressure from the
market information effect dominates the upward lift
from the operational mismatch effect, implying that
the firm’s most desired effort level \( e'(\theta) \) increases in \( \theta \)
on the regime of low values of \( \theta \). Contrast emerges
for the high values of \( \theta \), in which the result is
reversed. The firm’s most desired effort level \( e'(\theta) \)
decreases in \( \theta \) over the regime of high values of \( \theta \). To
induce the agent to adopt such effort strategy (i.e.,
first increasing and then decreasing in \( \theta \), the
firm needs to offer the agent marginal rewards that first
increase and then decrease in the sales volume. This
explains the result in Proposition 3 that the firm’s
optimal compensation scheme is S-shaped. Note that
in the procurement setting that is very different from
ours, Chu and Sappington (2009) show that the classi-
cal convexity optimality result from Laffont and Tirole
(1986) may not hold in the presence of the
supplier’s cost reduction effort. In contrast, the driver
we identify in our paper is the consideration of sup-
ply–demand mismatch costs.

Although the condition \( \Gamma''(z) \geq 0 \) is satisfied by the
news vendor type mismatch costs (almost every-
where), it is worth mentioning that the optimal com-
ensation plan \( T'(x) \) in the special case of piece-wise
linear mismatch cost, that is, \( \Gamma(z) = c_u z^+ + c_o (-z)^+ \)
with \( z = x - q \) (described in the model section), first
convex increases and then stays at constant as \( x \)
in creases when \( c_u \geq p \), and piece-wise convex
increases in \( x \) when \( c_u < p \) (the piece-wise feature is
caused by the fact that \( \Gamma(z) \) is nondifferentiable at \( z = 0 \)). Regardless, the convex increasing property of
the optimal compensation scheme without the mis-
match cost, as demonstrated in Proposition 2, no
longer holds even in this special case of piece-wise
linear mismatch cost.

The convexity assumption of \( H(\theta) \) is satisfied by the
uniform and exponential distributions. Further, we
numerically verify that the convexity assumption also
holds under the normal distributions with mean vary-
ing in \{1,1.5,2\} and standard deviation varying in
\{0.2,0.5,0.8\}, and the gamma distributions with shape
varying in \{1.5,2,2.5,3,3.5,4,4.5,5\} and scale varying in
\{0.5,1,5,2,2.5,3\}. Figure 1 depicts the firm’s optimal
compensation plan \( T'(x) \) for a representative numeri-
cal example where \( \Theta \sim N(4,1), \quad p = 10, \quad q = 5, \quad \Gamma(z) = z^2 \).

Proposition 3 implies that the consideration of oper-
ational costs in the design of the sales compensation
scheme has a clear qualitative impact on its structure: 
without operational considerations, the firm finds it
optimal to give increasing marginal rewards for sales
whereas when the firm takes operational costs into
account, she does not find such a scheme optimal any
more. Instead, the best sales compensation scheme
the firm should offer to the sales agent is guided by
the market information motivation of increasing sales
by offering an increasingly attractive scheme for the
agents (meaning increasing marginal rewards) up to a
certain threshold and beyond that, the compensation
scheme is guided by the operational motivation to
reduce the supply–demand mismatch.

Due to their ease of implementation, the piece-wise
linear compensation schemes with several pre-spec-
ified sales targets are frequently observed in practice,
under which marginal rewards may vary after meet-
ing a sales target. The theoretical optimality result in
Proposition 2 suggests that the firm should increase
the marginal rewards for meeting a higher sales target
without consideration of mismatch costs. In contrast,
with the consideration of mismatch costs, Proposition
3 suggests that the firm may need to lower the mar-
ginal rewards after reaching a certain sales target.
Interestingly, such S-shaped contracts are consistent
with the sales contracts illustrated in Oyer (2000).

As established by the agency literature on the opti-
mal compensation scheme (Laffont and Tirole 1986,
Rao 1990), the convex increasing property of the opti-
mal compensation scheme under the deterministic
demand case implies that the menu of linear compen-
sation (MLC) plans (with a higher commission rate
coupled with a lower fixed salary intended for a
higher type agent) is optimal for the more general
case with stochastic demand. However, the failure of
convexity property resulted from the consideration of
supply–demand mismatch costs implies that the opti-
mal compensation scheme for the stochastic demand
case needs not to be MLC. This is the issue we will
examine in the next section.

4. The Optimal Compensation
Scheme under Stochastic Demand

In this section, we relax the deterministic demand
assumption by imposing a random noise term in the
sales response function, that is, \( x = \Theta + e + \varepsilon \), where \( \varepsilon \)
is a zero-mean random variable with the standard
deviation denoted by \( \sigma \).

Unlike the deterministic demand case where the
direct mechanism consists of sales-payment pairs cor-
responding to the agent types, the firm can no longer
specify in the contract the exact sales volume due to
the random noise term \( \varepsilon \) that is beyond the agent’s
control. Instead, a general direct mechanism should
tie the agent’s compensation to his report of the mar-
ket condition and the realized sales. That is, the most
general form of a direct mechanism is \( t(x,\theta) \), under
which the agent’s compensation is \( t(x,\theta) \) if he reports
\( \theta \) and the realized sales volume is \( x \). With repetitive
Fig. 1 Optimal Compensation Plan [Color figure can be viewed at wileyonlinelibrary.com]

use of notation, let \( \pi(\theta, \hat{\theta}) \) denote the profit obtained by the type-\( \theta \) agent who reports \( \hat{\theta} \). Hence,

\[
\pi(\theta, \hat{\theta}) = \max_t E_t [t(\theta + e, \hat{\theta} - \hat{e}^2/2)].
\]

To ensure truth telling from the agent, we must have that for all \( \theta, \hat{\theta} \in [\bar{\theta}, \hat{\theta}] \),

\[
\pi(\theta, \hat{\theta}) \geq \pi(\theta, \hat{\theta}), \quad \text{(IC)}
\]

which indicates that it is in the best interest of the type-\( \theta \) agent to truthfully report \( \hat{\theta} \). Let \( e(\theta) \) be the type-\( \theta \) agent’s optimal effort level under truth telling. To ensure participation from the agent of any type, we must have that for all \( \theta \in [\bar{\theta}, \hat{\theta}] \),

\[
\pi(\theta, \theta) \geq 0. \quad \text{(IR)}
\]

Under the direct mechanism \( t(x, \theta) \) ensuring truth telling from the agent, the firm’s expected profits are \( E_{\theta, \epsilon} \{ \Pi(q, \Theta + e(\Theta) + \epsilon) - t(\Theta + e(\Theta) + \epsilon, \Theta) \} \). Therefore, the firm’s optimal direct mechanism is the solution to the following problem, denoted by (P’),

\[
\max_{t(x, \theta)} E_{\theta, \epsilon} \{ \Pi(q, \Theta + e(\Theta) + \epsilon) - t(\Theta + e(\Theta) + \epsilon, \Theta) \}
\]

s.t. IC’ and IR’.

Similar to Equation (1), we define

\[
e^*(\theta) = \arg \max_{e} \{ p e - e^2/2 - E_{\theta, \epsilon}(e + \theta + e - q) - H(\theta)e \}. \quad \text{(2)}
\]

Define

\[
\begin{align*}
\alpha(\theta) &= \int_{0}^{\theta} e^*(z) dz + [e^*(\theta)]^2/2 + \gamma \sigma^2 - e^*(\theta)(e^*(\theta) + \theta), \\
\beta(\theta) &= e^*(\theta), \\
\gamma &= \min{\bar{\gamma} \geq 0 | \frac{d\alpha^*(\theta)}{d\theta} + \bar{\gamma} \frac{d\alpha^*(\theta) + \theta}{d\theta} \geq 0 \text{ for all } \theta \in [\bar{\theta}, \hat{\theta}]}.
\end{align*}
\]

Note that \( \gamma \) is well defined because it is verifiable that \( d[e^*(\theta) + \theta]/d\theta > 0 \) and \( de^*(\theta)/d\theta > -1 \) for all \( \theta \in [\bar{\theta}, \hat{\theta}] \). Now we are ready to present the main result of this section.

**Proposition 4.** The firm’s optimal compensation scheme \( t^*(x, \theta) \) to the problem (P’) is:

\[
t^*(x, \theta) = \alpha(\theta) + \beta(\theta)x - \gamma(x - e^*(\theta) - \theta)^2,
\]

under which the type-\( \theta \) agent reports his true market condition \( \theta \), exerts effort \( e^*(\theta) \), and receives the compensation \( t^*(x, \theta) \) consisting of a fixed salary \( \alpha(\theta) \), commissions \( \beta(\theta)x \) that are linear in sales \( x \), and a quadratic penalty term for his forecast error \( x - e^*(\theta) - \theta \).

Recall that without considering the mismatch costs, that is, \( \Gamma(\cdot) \equiv 0 \), the sole presence of market information effect results in the monotonic increasing property of \( e^*(\theta) \). By definition of \( \gamma \), we have \( \gamma = 0 \). This, together with Proposition 4, implies...
that the firm’s optimal compensation scheme consists of a menu of linear contracts (MLC), a result that has been established by the extant literature on salesforce compensation.

Contrast emerges with the consideration of mismatch costs. Proposition 3 demonstrates that the firm’s most desired effort level from the type-\(\theta\) agent, that is, \(e^*(\theta)\), may decrease in \(\theta\) with the consideration of mismatch costs. To incentivize the agent to exert the firm’s most desired effort level, the commission rate tailored to each type agent must be aligned with the firm’s most desired effort level. The non-monotonicity of \(e^*(\theta)\) implies that the firm may have to offer a lower commission rate to a higher type agent relative to a lower type agent for effort incentive provision. Clearly, this is in conflict with the firm’s goal of inducing truth telling because the higher type agent has incentive to deviate to the contract intended for the lower type agent because the latter has a higher commission rate. Therefore, when the firm’s most desired effort level \(e^*(\theta)\) is not monotone increasing in \(\theta\), MLC suffers from the conflicts between the effort incentive provision and the truth telling incentive provision. The classical result of optimality of MLC breaks with the consideration of mismatch costs.

Interestingly, Proposition 4 reveals a simple way to modify the MLC to cope with the operational mismatch costs: Adding a forecast penalty term in MLC can restore optimality by eliminating the conflicts between the effort incentive provision and the truth telling incentive provision even if \(e^*(\theta)\) is not monotone increasing in \(\theta\). The intuition is that the firm can use the linear commission part solely to motivate the agent’s effort provision, because the forecast penalty does not interfere with effort incentive provision. Although the non-monotone commission rates \(\beta(\theta)\) may incentivize a higher-type agent to mimic a lower-type agent due to the higher commission rate intended for the lower-type agent, the forecast penalty with a sufficiently large penalty parameter \(\gamma\) can completely deter such behavior and ensure truth telling.

The optimal compensation scheme in Proposition 4 is reminiscent of the Gonik (1978) scheme, which was implemented by IBM’s Brazilian unit to extract market demand information and to motivate sales effort from the salesforce. Turner et al. (2007) describe that such forecast-based incentive schemes are being widely used in the pharmaceutical industry in Europe. Proposition 4 provides theoretical justification for the use of these forecast-based incentive schemes, especially in business environments where the supply–demand mismatch costs are significant.

5. Conclusions

By considering operational factors such as the supply–demand mismatch costs, we show that the classical results on the optimal compensation scheme may break, that is, the optimality of convex increasing compensation scheme for the case of deterministic demand response, and the optimality of MLC for the case of stochastic demand response. Instead, we show that for the former case, the optimal compensation scheme is S-shaped under certain conditions, and for the latter case, the optimal compensation scheme is MLC coupled with forecast penalties. Our results provide theoretical support for the practical use of S-shaped compensation schemes and the forecast-based incentive schemes, especially when the firm’s supply–demand mismatch costs are significant.

We remark that in our model, the firm’s inventory decision must be made before the agent observing the market condition. Such a sequence is applicable in business settings such as fashion/apparel and consumer electronics where firms often have to make inventory stocking decisions long before the selling season due to a relatively long production/order lead time. Under such a sequence, when the firm makes its inventory decision, the market information is often not yet available even to the agent. An alternative sequence of events is that the firm’s inventory decision can be made after the agent observing the market condition (see, e.g., Chen 2005). Intuitively, the firm can use the screened information via direct mechanisms to make an informed inventory decision so that the firm’s costs of supply–demand mismatch have no impacts on its compensation design problem, implying that consistent with the classical results, the firm’s optimal compensation scheme is convex under deterministic demand and MLC under stochastic demand.

Acknowledgments

The authors are grateful to the Department Editor, the Senior Editor, and two referees for their constructive comments. The first author, Binqing Xiao, has benefited from the support of the National Natural Science Foundation of China (U1811462, 71671083, 71720107001), and the Fundamental Research Funds for the Central Universities (011814380036).

Appendix

PROOF OF PROPOSITION 1. Let \(\pi(\theta) = \pi(\theta, \theta)\). The (IC) constraint, together with the Envelope Theorem,
implies that \( \frac{d\pi(\theta)}{d\theta} = \partial \pi(\theta)/\partial \theta |_{\theta = \theta} = x(\theta) - \theta \).

Taking the integral, together with the fact that the (IR) constraint must bind at \( \theta = \hat{\theta} \), that is, \( \pi(\theta) = 0 \), we have \( \pi(\theta) = \int_0^\theta (x(z) - z)dz \). This, together with the definition of \( \pi(\theta) \) that \( \pi(\theta) = t(\theta) - (x(\theta) - \theta)^2/2 \), allows us to write \( t(\theta) \) as a function of \( x(\theta) \):

\[
t(\theta) = (x(\theta) - \theta)^2/2 + \int_0^\theta (x(z) - z)dz. \tag{A1}
\]

Substituting \( t(\theta) \) with the right hand side of the above equation, we can rewrite the firm’s expected profits as

\[
E_\theta[\Pi(q, x(\Theta)) - t(\Theta)] = \int_0^\theta \{px(\theta) - \Gamma(x(\theta) - q) - (x(\theta) - \theta)^2/2 \}
- \int_0^\theta (x(z) - z)dz)f(\theta)d\theta
= \int_0^\theta \{px(\theta) - \Gamma(x(\theta) - q) - (x(\theta) - \theta)^2/2 \}
- \Gamma'(x(\theta))H(\theta)f(\theta)d\theta.
\]

Because we have assumed that \( \Gamma(\cdot) \) is convex, the whole function inside the integral is concave in \( x(\theta) \) for any given \( \theta \). We can then use point-wise optimization to obtain the optimal value of \( x(\theta) \), denoted by \( x^*(\theta) \), which maximizes the firm’s expected profits. Because the terms inside the above integral has a strict increasing difference in \( x(\theta) \) and \( \theta \), \( x^*(\theta) \) strictly increases in \( \theta \). Further, it follows from the definition of \( e^*(\theta) \) in Equation (1) that \( x^*(\theta) = e^*(\theta) + \theta \).

Substituting \( x(\theta) \) with \( x^*(\theta) \) in Equation (A1), we have

\[
t^*(\theta) = [e^*(\theta)]^2/2 + \int_0^\theta e^*(z)dz.
\]

Note that the direct mechanism \((x^*(\theta), t^*(\theta))\) maximizes the firm’s expected profits and satisfies the (IR) constraint. It remains to prove that it also satisfies the (IC) constraint. Under the direct mechanism \((x^*(\theta), t^*(\theta))\), we have

\[
\partial \pi(\theta)/\partial \theta = (\theta - \hat{\theta})dx^*(\hat{\theta})/d\theta,
\]

which, together with the fact that \( x^*(\cdot) \) is an increasing function, implies that the (IC) constraint is satisfied.

**Proof of Proposition 2.** The first statement of the proposition follows from the definition of \( e^*(\theta) \) in Equation (1) and the fact that \( pe - e^2/2 - H(\theta)e \) has an increasing difference in \( \theta \) and \( e \) due to the assumption that \( H(\theta) \) is a decreasing function. The last statement of the proposition then follows from the fact that \( \partial^2T(x)/\partial x^2 \) has the same sign as that of \( de^*(\theta)/d\theta \) at \( \theta = 0^+(x) \).

**Proof of Proposition 3.** It follows from the definition of \( e^*(\theta) \) in Equation (1) that the sign of \( de^*(\theta)/d\theta \) is the same as that of \( -\Gamma''(e^*(\theta) + \theta - q) - H'(\theta) \), which changes its sign from positive to negative at most once as \( \theta \) increases because its derivative \(-\Gamma''(e^*(\theta) + \theta - q) - H'(\theta) \leq 0 \) (following from the assumptions \( \Gamma''(z) \geq 0 \) and \( H'(\theta) \geq 0 \), and \( |e'(\theta) + \theta'| > 0 \), a result established in the proof of Proposition 1. Therefore, as \( \theta \) increases, \( e^*(\theta) \) first increases and then decreases, implying that \( T^*(x) \) is first convex increasing and then concave increasing in \( x \).

**Proof of Proposition 4.** We first use the first-order necessary condition of (IC') to derive an upper bound for the firm’s optimal expected profits under the problem (P'). We then show that the direct mechanism \( t(x, \theta) \) defined in Proposition 4 achieves this upper bound for the firm and satisfies (IC') and (IR') constraints, and thus is an optimal solution to (P').

Take any direct mechanism \( t(x, \theta) \). Let \( e(\theta) \) be the type-\( \theta \) agent’s optimal effort level under the direct mechanism \( t(x, \theta) \). Taking the partial derivative of \( \pi(0, \theta) \) with respect to \( \theta \), we have

\[
\partial \pi(0, \theta)/\partial \theta = E_t t_1(\theta + e(\theta, \theta) + \epsilon, \theta) \]

\[
= \partial \pi(0, \theta)/\partial \theta = e(0, \theta), \]

where \( t_1(\cdot, \cdot) \) is the partial derivative of \( x(\cdot, \cdot) \) with respect to the first parameter and \( e(\theta, \theta) \) is the type-\( \theta \) agent’s optimal effort level when reporting \( \theta \). This, together with the first-order necessary condition for the maximizer \( e(0, \theta) \), that is, \( E_t t_1(\theta + e(0, \theta) + \epsilon, \theta) = e(0, \theta) \), leads to \( \partial \pi(0, \theta)/\partial \theta = e(0, \theta) \). Letting \( \theta = \hat{\theta} \), we have \( \partial \pi(0, \theta)/\partial \theta |_{\theta = \hat{\theta}} = e(0, \hat{\theta}) \). It follows from the (IC') constraint and the Envelope Theorem that \( d\pi(0)/d\theta = \partial \pi(0, \theta)/\partial \theta |_{\theta = 0} \). Hence, \( d\pi(0)/d\theta = e(0) \).

Taking the integral, together with the (IR') constraint that \( \pi(\theta) = 0 \), we have

\[
E_t t(\theta + e(\theta) + \epsilon, \theta) = [e(\theta)]^2/2 + \int_0^\theta e(z)dz.
\]

Consequently, we can rewrite the firm’s expected profits as
Because we have assumed that \( \Gamma(\cdot) \) is convex, the whole function inside the expectation is concave in \( e(\theta) \) for any given \( \theta = \theta \). We can then use point-wise optimization to obtain the optimal value of \( e(\theta) \), denoted by \( e'(\theta) \), which maximizes the firm’s expected profits. Note that \( e'(\theta) \) is the same as that defined in Equation (2). Substituting \( e(\theta) = e'(\theta) \) for every \( \theta \) into the above equation, we obtain an upper bound on the firm’s optimal expected profits under the problem (P).

It remains to show that the direct mechanism \( t^*(x, \theta) \) defined in Proposition 4 achieves this upper bound for the firm, satisfies (IC') and (IR') constraints, and induces the optimal effort \( e'(\theta) \) from the type-\( \theta \) agent.

Under the constructed compensation plan \( t^*(x, \theta) \), we have

\[
\pi(\theta, \hat{\theta}) = \max_{\epsilon} \left( \pi(\hat{\theta}) + \beta(\hat{\theta})(\theta + e) - \gamma E_e(\theta + e + e'(\hat{\theta}) - \hat{\theta})^2 - e^2/2 \right) \\
= \max_{\Delta} \left( \pi(\hat{\theta}) + \beta(\hat{\theta})\Delta - \gamma E_e(\Delta + e - e'(\hat{\theta}) - \hat{\theta})^2 - (\Delta - \theta)^2/2 \right),
\]

where \( \Delta = \theta + e \). Clearly, the function inside the above maximization has an increasing difference in \( \Delta \) and \( \theta \), implying that the optimal \( \Delta \) increases in \( \theta \), or equivalently, \( \theta + e(\theta, \hat{\theta}) \) increases in \( \theta \) where \( e(\theta, \hat{\theta}) \) is the type-\( \theta \) agent’s optimal effort when he reports \( \hat{\theta} \).

Taking the partial derivative with respect to \( \hat{\theta} \), we have

\[
\partial \pi(\theta, \hat{\theta})/\partial \hat{\theta} = \left( [e'(\hat{\theta})]' + \gamma [e'(\hat{\theta}) + \hat{\theta}]' \right) (\theta + e(\theta, \hat{\theta}) - e'(\hat{\theta}) - \hat{\theta}).
\]

Note that we have proved in the previous paragraph that \( \theta + e(\theta, \hat{\theta}) \) increases in \( \theta \). This, together with the definition that \( e(\theta, \hat{\theta}) = e'(\theta) \), implies that the second part on the right hand side of the above equation, that is, \( \theta + e(\theta, \hat{\theta}) - e'(\hat{\theta}) - \hat{\theta} \), is non-negative for \( \hat{\theta} < \theta \) and non-positive for \( \hat{\theta} > \theta \). Therefore, to ensure that (IC') is satisfied, it suffices to set \( \gamma \) such that the first term on the right hand side of the above equation, i.e., \( [e'(\hat{\theta})]' + \gamma [e'(\hat{\theta}) + \hat{\theta}]' \), is positive. This leads to the following definition of \( \gamma \) that satisfies (IC'):

\[
\gamma = \min \{ \gamma \geq 0 : \partial e(\theta)/\partial \theta + \gamma\partial e'(\theta)/\partial \theta \geq 0 \text{ for all } \theta \in [\hat{\theta}, \theta] \}.
\]

Under truth telling, the type-\( \theta \) agent’s expected profit at any given effort level \( e \) is equal to

\[
\pi(\theta) + e'(\theta)(\theta + e) - \gamma E_e(\theta + e - e'(\theta) - \theta)^2 - e^2/2,
\]

which is concave in \( e \) and maximized at \( e = \beta(\theta) = e'(\theta) \) (by the first-order condition and the definition of \( \beta(\theta) \)). Therefore, it is optimal for the type-\( \theta \) agent to truthfully report \( \theta \) and exerts the effort level \( e'(\theta) \). This, together with the definition of \( \pi(\theta) \), implies that that the type-\( \theta \) agent’s optimal expected profits can be written as

\[
\pi(\theta) + e'(\theta)(\theta + e) - \gamma E_e(\theta + e - e'(\theta) - \theta)^2/2
= \int_0^\theta e'(z)dz + [e'(\theta)^2 + \gamma\sigma_e^2 - e'(\theta)(e'(\theta) + \theta)]/2
+ e'(\theta)(\theta + e') - \gamma E_e(\theta + e')^2/2
= \int_0^\theta e'(z)dz,
\]

implying that (IR') is satisfied for any \( \theta \).

Under the agent’s optimal response, the firm achieves the upper bound in her expected profits and hence the compensation scheme \( t^*(x, \theta) \) is optimal.

References


