# FORECASTING INCENTIVES BASED ON VALUE MARGINAL PRODUCT

Carl Lundgren
VALMARPRO FORECASTING, INC.
2121 Columbia Pike, Suite 301
Arlington, VA 22204-4423

(703) 521-1037

Version Vb

June 1995

The author would like to thank Kent Osband for useful comments on a previous version of this paper.

## NOTICE OF PATENT ISSUED

U.S. Patent 5,608,620.

This paper describes a method of economic incentives, involving plural-forecaster payment systems, upon which the author and inventor has been issued a patent. (U.S. Patent 5,608,620) The patent on this invention only restricts actual use of the described invention; it does not restrict in any way the verbal or written discussion, description, or criticism of that invention. Any actual use of the patented invention without the express written permission of Valmarpro Patents, Inc. is strictly prohibited.

ABSTRACT

FORECASTING INCENTIVES BASED ON VALUE MARGINAL PRODUCT

This paper describes a scheme of cost-efficient incentives for eliciting unbiased predictions from human forecasters. The method measures a close proxy for the value marginal product (VMP) of each forecaster and pays in accordance therewith. The payment method results in optimal exertions of effort by forecasters, and attracts nearly optimal numbers of forecasters to perform the forecasting task. The method works very well when forecasters are risk neutral, but may introduce some bias when forecasters are risk averse. Methods for dealing with the potential bias from risk aversion are discussed.



J.E.L. Codes: D80 Information and Uncertainty
D84 Expectations; Speculations
G14 Information and Market Efficiency

Keywords: forecasting, incentives, VMP, predictions.

**Introduction**

The economic need for accurate forecasting of future events, whether they be prices, quantities, values, or other variables, is pervasive. Economic variables of interest to both public and private policy makers include future prices and quantities of commodities, expected future profits of business firms, and expected quantities, damages, and marginal costs of pollutants. The purpose of this paper is to describe the design of an efficient scheme of incentives for eliciting unbiased predictions from two or more human forecasters.[1] The approach is to find a good proxy measure for the value marginal product (VMP) of a forecaster's predictions, so that each forecaster may be paid in accordance with his contribution to a collective forecast. The corresponding incentive scheme applied to only one forecaster would not, in general, be unbiased.

Osband (1989) also attempts to derive optimal forecasting incentives. The Osband method differs, in that only one forecaster is hired *ex post*, whereas the present method hires at least two forecasters. This difference is an important advantage, since it allows comparisons between forecaster predictions *ex post*. Because of these comparisons, the present method allows a reduction in the variance of compensation which risk-averse forecasters might otherwise need to suffer. Another advantage of considering incentive schemes with multiple forecasters is that it is frequently imprudent (due to incomplete information or biased judgement of a single forecaster) or too costly (due to increasing marginal cost of effort) to rely on the predictions of only one forecaster.

Osband's thesis (1985) is mainly devoted to analyzing one-forecaster incentive schemes, but section 5.5 does discuss a scheme with multiple forecasters. Osband (1985) derives an "optimal" incentive scheme assuming the principal's loss function is quadratic, and assuming an arbitrary restriction on the allowable payoff structure.[2] After adjusting for considerable differences in mathematical notation, it can be shown that the incentive scheme derived by Osband (pp. 113-114) is similar (not identical) to the incentive scheme of this paper, after plugging in the assumption of quadratic loss functions. No reference is made by Osband to any economic intuition concerning VMP, nor is it clear how his techniques might be generalized to nonquadratic loss functions.

Two other papers, Kadane and Winkler (1988) and Page (1988), also suggest incentive schemes for eliciting forecasts, but are more narrowly focused on the prediction of probabilities, rather than events or values in general.

Section I discusses the general problem of setting up an incentive scheme to motivate forecasters. Section II suggests using the VMP concept as an intuitive way of developing pay schedules for forecasters. Section III tests the candidate pay schedule and presents two propositions. Section IV suggests that forecaster risk aversion and other considerations will cause forecasters to form firms and partnerships, and that this market outcome is the most practical way of handling forecasting risk. Section V analyzes an example to show that forecaster efforts are optimally determined. Section VI shows that the number of forecasters attracted to the forecasting task tends to be very close to optimal. Section VII concludes.

## I. Statement of the Problem

Suppose that the goal of the principal (forecast requisitioner) is to obtain an accurate prediction concerning the future realization of a random variable X. Suppose further, that this goal is to be accomplished indirectly, rather than directly, by hiring a set of agents (forecasters) who will do the actual forecasting. The problem for the forecast requisitioner is to find a set of contracts for the forecasters such that the incentives given to the forecasters result in tolerably good forecasts at a tolerably low cost. We further suppose that the principal is unsophisticated, and cannot condition the parameters of the incentive contracts on any detailed knowledge of how the forecasters perform their task.

When forecasters' predictions differ, there is a need to aggregate individual predictions to obtain a collective prediction suitable for further action. A typical method of aggregation might be to take an average or weighted average of forecasters' predictions, such as an arithmetic mean or a geometric mean. Let $X_c$ represent the vector of individual predictions, $X_1$, $X_2$, ..., $X_n$ of forecasters 1, 2, ..., n. Suppose that we have a well-defined prediction aggregator function which yields specific collective predictions when given information concerning any one or more predictions from individual forecasters. Such a prediction aggregator function might be generalized as follows:

$$G(X_c) = G(X_1, X_2, X_3, ..., X_n) \qquad (1)$$

The function G is assumed to be well-defined for any number of variables ($n \geq 1$), so that expansion or contraction of the variable set will still provide a well-defined answer. Let $X_{ci}$ represent the vector of predictions of all forecasters except forecaster i. If the set $X_c$ contained at least two predictions, then the vector $X_{ci}$ is well-defined. $G(X_{ci})$ is a "secondary collective prediction," which would presumably be issued in the absence of forecaster i's prediction.

Let $B(X_a, G(X_c))$ be the benefits which accrue when $G(X_c)$ is the collective prediction of X, while $X_a$ is an actual or estimated value of X which is later observed. The loss function, $L(X_a, G(X_c))$, tells us the lost benefits which occur when the predicted X differs from its actual value:

$$L(X_a, G(X_c)) = B(X_a, X_a) - B(X_a, G(X_c)) \qquad (2)$$

The value $X_a$ can be used as a "criterion value"--a variable value which is used to judge the accuracy or inaccuracy of forecasters' predictions. If the actual value of the variable being predicted is observed within a reasonable period of time, it is natural to use the actual variable value as the criterion value. Otherwise, it will be necessary to use a proxy.[3]

The goal of society is to minimize the sum of a) the welfare loss from erroneous prediction, $L(X_a,G(X_c))$, plus b) the opportunity costs of forecaster effort, plus c) the costs of the risk premia needed to compensate risk-averse forecasters for their acceptance of risk. Section II argues that a payment scheme based on VMP would automatically balance considerations a) and b) internally to the forecaster. Section IV argues that the problem of insuring forecaster risk is best dealt with as a separate transaction. Hence, as an initial simplification, assume that the opportunity costs of forecaster effort are already sunk and that society's goal is simply to elicit unbiased predictions, given the information sets already at forecasters' disposal.

To accomplish this task, society must choose a payment schedule (P) for each forecaster such that each forecaster is motivated to provide a prediction which minimizes the expected loss, E(L). The payment schedule for each forecaster can be made a function of $X_a$ and each $X_i$: $P_i=P_i(X_a,X_1,X_2,X_3,...,X_n)$ $=P_i(X_a,X_i,X_{ci})$. Given the payment schedule, each forecaster will choose his prediction to maximize his own utility, given his own utility function which we may presume is not directly observed by others.

Let $f(X_a)$ be a probability density function which is based on the combined information sets of all forecasters.[4] The optimal collective prediction (G*) minimizes the expected loss:

$$E(L) = \int_{-\infty}^{\infty} f(X_a)L(X_a,G^*)\partial X_a \qquad (3)$$

Choosing G* to minimize E(L) means that the following first-order condition must be satisfied:

$$\frac{\partial}{\partial G^*} \int_{-\infty}^{\infty} f(X_a)L(X_a,G^*)\partial X_a = 0 \qquad (4)$$

G* is not directly observable or computable by the forecast requisitioner, since the information sets on which G* is based are not directly available to the requisitioner. Instead, it is necessary for the requisitioner to choose a payment scheme which motivates each forecaster to choose individual predictions such that the collective prediction, $G(X_c)$, tends to satisfy the above condition for G*.

## II. The VMP Method of Solution

The approach of this paper is to choose a set of contracts which economic intuition suggests is likely to have good properties, to investigate those properties, and to try to evaluate whether the set of contracts so chosen has sufficiently good properties. The alternative procedure, attempting to find a "best" set of contracts as a solution to some sophisticated optimization problem, is unlikely to yield useful answers when, as here, the problem a) is stated with great generality, b) involves multiple agents, c) is subject to a fuzzy constraint of economic practicability due to unsophistication of the principle, and d) the relevant mathematical optimization technique is known to be extremely complex even for simple problems.

The guiding economic intuition used here is that paying forecasters according to their value marginal product (VMP) is likely to have good incentive effects in terms of both attracting the right number of forecasters and motivating the right level of effort. If we can accurately measure both the cost and the expected VMP of each forecaster, then we can hire forecasters until the cost of an additional forecaster equals his VMP. This would assure forecasting efficiency on the extensive margin (optimal number of forecasters). Additionally, if we can observe the VMP of each forecaster, we can compensate each forecaster in accordance therewith. This would assure forecasting efficiency on the intensive margin (optimal intensity of effort per forecaster). Despite the unobservability of mental effort, compensation according to forecaster VMP assures that each forecaster will continue to exert mental efforts until the marginal cost of an extra unit of mental effort equals its marginal benefit in terms of its expected increase in VMP.

Payment according to VMP requires some definition and measurement of VMP in the field of forecasting.[5] The proxy for VMP used here is designated the "marginal contribution."[6] The marginal contribution asks how the value of a collective forecast changes, when the prediction of a particular forecaster is either contributed or withheld. The marginal contribution of forecaster i towards the accuracy of the collective forecast can be given by the equation:

$$MC_i = B(X_a,G(X_c))-B(X_a,G(X_{ci}))$$
$$= L(X_a,G(X_{ci}))-L(X_a,G(X_c)) \quad (5)$$

The marginal contribution for a particular forecaster might well be positive, zero, or negative, depending on whether $X_i$ moves the collective forecast towards or away from $X_a$. Typically, the sum of the marginal contributions for all forecasters combined will be positive. Typically, also, the expected marginal contribution (before observation of $X_a$) of each forecaster would be positive, as well, if we assume that each forecaster has at least some information of value to contribute to the collective prediction.

Hence, a natural candidate for the pay schedule of each forecaster would look something as follows:

$$P_i(X_i,X_{ci},X_a) = F + kL(X_a,G(X_{ci})) - kL(X_a,G(X_i,X_{ci})), \quad (6)$$
$$\text{where } k>0.$$

In the above equation, $G(X_c)$ is written out as $G(X_i,X_{ci})$ so as to emphasize that forecaster i's payment is contingent both on his own prediction ($X_i$) and the predictions of others ($X_{ci}$). The payment schedule in (6), of course, is simply a constant multiple of the VMP formula in equation (5). It remains only to test whether this payment schedule accomplishes its intended purpose.

## III. Properties of the VMP Method

Two propositions about VMP forecasting incentives can be stated:

<u>Proposition 1:</u> When all forecasters have identical beliefs and information sets concerning the probability distribution of

$X_a$, the incentive scheme in (6), combined with an optimal prediction aggregator function, yields optimal individual and collective predictions, regardless of whether forecasters are risk neutral or risk averse, provided at least two forecasters issue predictions.

<u>Proposition 2:</u> When all forecasters are risk neutral, the incentive scheme in (6), combined with an optimal prediction aggregator function, yields optimal collective predictions, regardless of whether or not forecasters have identical beliefs or information sets about the probability distribution of $X_a$.

We first define an optimal prediction aggregator function. By the conditions of the problem, each forecaster i is constrained to base his own forecast $X_i$ on his own information set $I_i$,[7] so that we may posit the existence of functions $X_i=X_i(I_i)$ and $X_{ci}=X_{ci}(I_{ci})$. An optimal prediction aggregator function is a function G which, given that the predictions of each forecaster are reported in accordance with the function $X_i(I_i)$,[8] chooses the optimal collective prediction G*, given the combined information sets of all forecasters.[9]

$$G = G(X_1(I_1),X_2(I_2),X_3(I_3),...,X_n(I_n)) \quad (7)$$
$$= G^*(I_1,I_2,I_3,...,I_n)$$

We may assume that each forecaster, indexed by i, has a utility function in wealth (or income) of $U_i(W)$. For Proposition 1, the forecaster must choose $X_i$ to maximize his expected utility under the payment scheme, where the expected utility is:[10]

$$\int_{-\infty}^{\infty} U_i(W_i+P_i(X_i,X_{ci},X_a))f(X_a)\partial X_a \quad (8)$$

$$= \int_{-\infty}^{\infty} U_i[W_i+F+kL(X_a,G(X_{ci}))-kL(X_a,G(X_i,X_{ci}))]f(X_a)\partial X_a \quad (9)$$

Solving the above problem requires taking partial derivatives with respect to $X_i$ and setting them equal to zero:

$$\int_{-\infty}^{\infty} U_i'[W_i+F+kL(X_a,G(X_{ci}))-kL(X_a,G(X_i,X_{ci}))] *[-k(\partial L/\partial G)(\partial G/\partial X_i)]f(X_a)\partial X_a = 0 \quad (10)$$

For analysis of Proposition 1, let G* be the optimal prediction from equation (4). All forecasters are agreed that G* is the optimal prediction. Assume further that the forecast requisitioner has chosen an aggregator function such that if all forecasters choose G*, then G* is the collective prediction.[11] Hence, if all other forecasters choose G* as their prediction, then $G(X_{ci})=G^*$. If $X_i=G^*$ properly solves the equation under these circumstances, then each forecaster is properly motivated and we have a Nash equilibrium where each forecaster submits the optimal prediction.

If we substitute $X_i=G(X_{ci})=G(Xi,Xci)=G^*$ into (10), the argument of Ui' becomes a constant, since $[L(X_a,G^*)-L(X_a,G^*)]=0$. The first-order condition in (10) then reduces to:

$$U_i'[W_i+F](-k)(\partial G/\partial X_i) \int_{-\infty}^{\infty} (\partial L/\partial G)f(X_a)\partial X_a = 0 \quad (11)$$

From equation (4), this integral equals zero when $X_i=G(X_{ci})=G^*$. Hence, it is a Nash equilibrium for all forecasters to choose $X_i=G^*$, even if risk-averse.[12] Another way of seeing this is that a risk-averse forecaster wants to forecast truthfully, because the truthful report receives a constant payoff of F while the non-truthful report receives a random payoff with expected value less than F. At least when forecasters are agreed concerning the probability distribution of $X_a$, risk aversion does not bias the forecaster's prediction and the actual extent of risk aversion is irrelevant to the optimal functioning of this forecasting method.

For the more realistic situation of Proposition 2, assume that different forecasters have different opinions about the probability distribution for the variable X. This leaves open the possibility that $X_i$ differs from $G(X_{ci})$, and that a forecaster may have advance awareness of this fact. Suppose that forecaster predictions differ because they have access to different (but possibly overlapping) information sets. Each forecaster i, after observing information $I_i$, must choose $X_i$ to maximize his expected utility in the following double integral:[13]

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U_i[W_i+F+kL(X_a,G(X_{ci}(I_{ci})))-kL(X_a,G(X_i,X_{ci}(I_{ci})))] * f(X_a|I_i,I_{ci})g(I_{ci}|I_i)\,\partial X_a\,\partial I_{ci} \qquad (12)$$

We may now ask the question of whether it is optimal for forecaster i to issue predictions according to the function $X_i(I_i)$ if he assumes that all other forecasters j are issuing their predictions according to the functions $X_j(I_j)$. If the answer is yes, the scheme is incentive compatible. If the answer is no, forecasters have a moral-hazard temptation to issue biased predictions.

Solving the above problem requires taking partial derivatives with respect to $X_i$ and setting them equal to zero:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} U_i'[W_i+F+kL(X_a,G(X_{ci}(I_{ci})))-kL(X_a,G(X_i,X_{ci}(I_{ci})))] * [-k(\partial L/\partial G)(\partial G/\partial X_i)]f(X_a|I_i,I_{ci})g(I_{ci}|I_i)\ \partial X_a\ \partial I_{ci} = 0 \quad (13)$$

We may split the above double integral into two double integrals, using the following additive identity:

$$U_i'[W_i+F+kL_{ci}-kL_c] = U_i'[W_i+F] + \{U_i'[W_i+F+kL_{ci}-kL_c]-U_i'[W_i+F]\} \quad (14)$$

This yields the following:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} U_i'[W_i+F] * [-k(\partial L/\partial G)(\partial G/\partial X_i)]f(X_a|I_i,I_{ci})g(I_{ci}|I_i)\ \partial X_a\ \partial I_{ci} \quad (15)$$

$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \{U_i'[W_i+F+kL_{ci}-kL_c] - U_i'[W_i+F]\} * [-k(\partial L/\partial G)(\partial G/\partial X_i)]f(X_a|I_i,I_{ci})g(I_{ci}|I_i)\ \partial X_a\ \partial I_{ci} = 0$$

Suppose now that forecaster i chooses $X_i$ according to $X_i(I_i)$ so that $G(X_i,X_{ci})=G^*$ as indicated in (7). We must now ascertain whether this choice of $X_i$ solves (15). Using similar reasoning as was used with equations (9) and (10), the first double integral in (15) vanishes when $G(X_i,X_{ci})=G^*$. However, the second double integral is not necessarily zero in general. The second double integral will be zero if the marginal utility, $U_i'$, is constant over the relevant range of wealth. Hence, we are completely assured of incentive compatibility for this scheme only if marginal utility is constant, meaning that forecasters must be risk neutral.[14]

## IV. Risk, Bias, and Intermediaries

If forecasters are risk averse, they will have a tendency to want to bias their predictions toward the expected or perceived value of $G(X_{ci})$ (which minimizes risk if $G(X_{ci})$ is known with certainty), rather than the socially optimal value of $X_i$ which would maximize expected forecaster payment. This bias cannot occur unless the forecaster knows, or can reasonably infer, the probable direction and magnitude of the difference between $X_i(I_i)$ and $G(X_{ci})$. When forecasts are offered simultaneously, $G(X_{ci})$ does not become known until after $X_i$ has already been submitted, so that the optimal bias may be (near) zero, because the forecaster's ignorant best guess of $G(X_{ci})$ is that its expected value is near $X_i(I_i)$.

It is not at all certain that this bias, if it occurs, would be serious. If it should prove serious, there are several ways of dealing with the potential bias of risk-averse forecasters: One way is to alter the payment scheme slightly so as to render forecasters effectively risk neutral.[15] A second way is to reduce risks by making sure that the potential variance in compensation is rather low compared to forecaster wealth. This might be accomplished, either by setting $k<1$, or by hiring so many forecasters that the expected variance in VMP is small for any one forecaster. A third way is to adjust the collective prediction to compensate for presumed bias, before using the prediction for further practical purposes.[16] A fourth way, indicated below, is to consider more seriously the possible role of firms and partnerships as intermediaries between forecasters and the forecast requisitioner.

Suppose that there can exist risk-neutral intermediaries between forecasters and forecast requisitioners. An intermediary can perform two functions: insure the forecaster and deliver risk-neutral predictions to the requisitioner. If the intermediary can (imperfectly) monitor forecaster efforts, it will be possible to ameliorate the moral hazard problem so as to offer (partial) insurance to risk-averse forecasters.

Even without the ability to monitor forecaster efforts, the intermediary can obtain risk-neutral predictions from the forecaster in exchange for a promise to pay the forecaster in accordance with subsequent observation of $X_i$, $G(X_{ci})$, and $X_a$. That is, given a statement from the forecaster concerning the optimal risk-neutral $X_i$ and a statement of the forecaster's level of risk aversion and the expected variance of $X_a$ about $G(X_c)$, the intermediary can compute the value of $X_i$ which the forecaster would have desired to submit, had the forecaster known in advance the value of $G(X_{ci})$. The intermediary performs a valuable service to the forecaster, because the forecaster acting alone cannot condition his forecast on $G(X_{ci})$ prior to observing $G(X_{ci})$. The intermediary also performs a valuable service for the forecast requisitioner, because the intermediary is able to issue risk-neutral forecasts to the requisitioner, which cannot be obtained from risk-averse forecasters acting alone.

Performance of this intermediating role requires some sophistication concerning the details of how the forecasting task is performed, a level of sophistication which we have assumed the requisitioner probably lacks. This approach does raise the interesting question of which types of contracts would be optimal for intermediaries to offer to forecasters. Nevertheless, from the perspective of the forecast requisitioner, which this paper takes, it is sufficient for the requisitioner simply to act as if forecasters (or their chosen intermediaries) are risk neutral, and let market intermediaries provide for the insurance function.

## V. Intensive Margin, an Example

From section III we learned that risk-neutral forecasters make unbiased predictions under this incentive scheme. The verbal intuitions explained in Sections I and II suggest that if k=1 in equation (6) then the efforts of forecasters will be optimally determined, given the number of forecasters. We now verify this intuition using a specific example.

In this example, assume that forecasters are risk neutral, that all random variables are normally distributed, and that the loss function takes the quadratic form:

$$L(X_a,G(X_c)) = h(X_a-G(X_c))^2, \quad h>0 \tag{16}$$

Since the loss function is quadratic, the optimal prediction is the expected value of $X_a$. We set h=1, since it makes no difference to the results.

Suppose further that X is the sum of two random variables, a humanly observable signal, S, and an unpredictable component, $E_a$. Each forecaster observes $I_i$, which is an observation of S that is clouded by a forecaster-specific error term $E_i$. Each error term is independent of all other error terms and also of $E_a$ and S. The variables are defined or distributed as follows:

$$X_a = S + E_a$$
$$I_i = S + E_i$$
$$S \sim N(0,\sigma_s^2) \qquad (17)$$
$$E_a \sim N(0,\sigma_a^2)$$
$$E_i \sim N(0,1/\tau_i)$$

Perhaps due to differences in opportunity, effort, or skill, the expected precision ($\tau_i$) of each forecaster may well be different. In keeping with our assumption that the forecast requisitioner is unsophisticated, we assume a) that the requisitioner has no advance knowledge of the proper weights to be attributed to each forecast, and b) does not know how the various $I_i$'s should be aggregated to determine the optimal prediction, given the $I_i$'s. We assume, however, that the forecasters themselves have the necessary sophistication to perform both tasks, provided they are properly motivated. Suppose, therefore, that each forecaster submits a prediction, $X_i$, and an expected precision, $T_i$, and that the forecast requisitioner aggregates predictions in the following simple-minded way:

$$G(X_c) = \sum_{i=1}^{N} T_i X_i \; / \; T_c, \qquad (18)$$

where $T_c = \sum_{i=1}^{N} T_i$.

The requisitioner simply takes a weighted average of each prediction $X_i$, based on the submitted weights, $T_i$, of each forecaster. The optimal collective forecast is computed as follows:[17]

$$G^* = \sum_{i=1}^{N} \tau_i ßI_i \;/\; \tau_c, \qquad (19)$$

where $\tau_c = \sum_{i=1}^{N} \tau_i$ and $ß = \sigma_s^2/(\sigma_s^2+1/\tau_c)$.

Given the aggregator function in (18), it is sufficient for unbiasedness that $T_i=\tau_i$ and $X_i=ßI_i$ for all forecasters. Note that the optimal $X_i$ depends on $\tau_c$. Since $\tau_c$ is not known in advance by each forecaster (though each forecaster may have a fair idea of the likely range), each forecaster would prefer to make his forecast conditional on $T_c$. Hence, let each forecaster submit both the conditional function $X_i(T_c)$ and the unconditional weight $T_i$.

Suppose now that the forecast requisitioner provides the following definitions:

$$T_{ci} = \sum_{j\#i} T_j \qquad (20)$$

$$G(X_{ci}) = \sum_{j\#i} T_j X_j \;/\; T_{ci}$$

and sets up the following pay schedule:[18]

$$P_i(T_i, X_i(T_i), T_{ci}, X_{ci}(T_{ci}), X_a) \qquad (21)$$

$$= k(X_a-G(X_{ci}))^2 - k(X_a-G(X_c))^2, \qquad \text{where } k>0.$$

<u>Proposition 3:</u> If the random variables $X_a$ and $I_i$ are specified as in (17), the forecasts are aggregated according to (18) and (20), and forecasters wish to maximize their expected payoffs, where this payoff is given (for any $k>0$) by (21), then it is a Nash equilibrium for each forecaster to report a truthful

precision value ($T_i=\tau_i$) and a conditional prediction function, $X_i(T)$. Further, this vector of reports, when aggregated according to (18), will minimize the expected value of the loss function given in (16).

Proof: See Appendix A.

Proposition 4: Under the conditions of Proposition 3, if k=1 in the pay schedule in (21), then, given the number and precision levels of the other forecasters, each forecaster exerts the socially optimal level of effort.

Proof: If we substitute $T_i=\tau_i$ and $X_i=ß_c I_i$ into (A.3) and take the unconditional expectation, we obtain:

$$E(P) = \sigma_s^4/(\sigma_s^2+1/\tau_c) - \sigma_s^4/(\sigma_s^2+1/\tau_{ci}) \quad (22)$$

Since $\tau_c=\tau_i+\tau_{ci}$, E(P) is a function of $\tau_i$. Express this relationship as $P(\tau_i)$. There is also a cost of exerting effort, which results in a given level of precision. Express this relationship as $C(\tau_i)$. The risk-neutral forecaster must solve:

$$\text{maximize } U(\tau_i) = P(\tau_i) - C(\tau_i) \quad (23)$$

This has solution:

$$\partial U/\partial\tau_i = P'(\tau_i) - C'(\tau_i) = 0 \quad (24)$$

Taking derivatives of (22) while taking $\tau_{ci}$ as given, we obtain:

$$C'(\tau_i) = P'(\tau_i) = \sigma_s^4/(\sigma_s^2\tau_c+1)^2 \quad (25)$$

The social welfare problem (holding constant for the number and type of forecasters) requires that a forecaster set forth the following amount of effort:

$$\text{maximize } SW(\tau_i) = -L(\tau_{ci},\tau_i) - C(\tau_i) - C(\tau_{ci}) \quad (26)$$

This has solution:

$$\partial SW/\partial\tau_i = -L'(\tau_i) - C'(\tau_i) = 0 \qquad (27)$$

To compute $L'(\tau_i)$, we must first compute E(L). Breaking down the variables in (16) into their component parts, we have:

$$L = [S+E_a - ß_c(S+E_c)]^2 \qquad (28)$$

$$= [(1-ß_c)S + E_a - ß_cE_c]^2$$

Taking expectations:

$$E(L) = \sigma_s^2/(\sigma_s^2\tau_c+1)^2 + \sigma_a^2 + \sigma_s^4\tau_c/(\sigma_s^2\tau_c+1)^2 \qquad (29)$$

$$= \sigma_s^2/(\sigma_s^2\tau_c+1) + \sigma_a^2$$

Hence, substituting into (27) we derive:

$$C'(\tau_i) = -L'(\tau_i) = \sigma_s^4/(\sigma_s^2\tau_c+1)^2 \qquad (30)$$

Comparison of (25) and (30) shows that the forecaster always exerts the socially optimal level of effort. Hence, there is always efficiency on the intensive margin. We now turn to a discussion of the extensive margin.

## VI. Extensive Margin, Identical Forecasters

There are several questions which arise on the extensive margin. One is whether the optimal number of forecasters enter the market, in the sense of volunteering to perform the forecasting task. The second is whether the optimal type of forecaster (or optimal combination of types) enters the market. Forecasters can differ according to type in at least two ways: a) forecasters may differ according to the level of precision which they find optimal to perform, and b) forecasters may differ according to the level of cost per unit of precision. Finally, there is an integer constraint, since forecasters do not come in fractional units.

We simplify our analysis by assuming that all forecasters, if they enter the market, subsequently find it optimal to choose effort levels which result in the same level of precision. Suppose, therefore, that each forecaster receives information with identical precision, $\tau$. As the previous section showed, the choice of $\tau$ is always optimal, given the number of forecasters who enter the market. $\tau$ is a function of the number of forecasters who enter the market. Generally, the larger the number of forecasters, the greater the precision of the collective forecast, and the smaller the optimal effort level per forecaster.

Given that the number of forecasters (N) must be an integer, the optimal number of forecasters ($N_o$) must solve the following inequalities in N:

$$-L(N\tau(N))-NC(\tau(N)) \geq -L((N-1)\tau(N-1))-(N-1)C(\tau(N-1)) \quad (31)$$

$$-L(N\tau(N))-NC(\tau(N)) \geq -L((N+1)\tau(N+1))-(N+1)C(\tau(N+1))$$

The forecasting incentives described in this paper would yield an equilibrium number of forecasters ($N_e$) which solves the following inequalities in N:

$$-L(N\tau(N)) + L((N-1)\tau(N)) \geq C(\tau(N)) \quad (32)$$

$$-L((N+1)\tau(N+1)) + L(N\tau(N+1)) \geq C(\tau(N+1))$$

The above two sets of inequalities will be precisely identical only if $\tau(N)=\tau(N+1)=\tau(N-1)$. This can occur only if effort levels are unaffected by the changes in incentives which result from a change in N (effort is perfectly inelastic). Even when effort levels vary, it will sometimes happen that the same integer N solves both sets of inequalities. For example, if (31) requires $2.13 \leq N_o \leq 3.03$ and (32) requires $2.56 \leq N_e \leq 3.43$, it is evident that $N_o=N_e=3$ solves both sets of inequalities. In such cases, there is optimality on both the intensive and extensive margins, if we assume that only efficient forecasters enter the market.[19]

The other possible source of inefficiency might arise if costlier types of forecasters are able to enter the market, but are not induced to leave. Because the number of forecasters must be an integer, the profits of forecasters are not necessarily driven to zero. Since efficient forecasters can earn positive profits without inducing further entry (which would cause a discontinuous fall in average profits), it is therefore possible for forecasters with somewhat higher costs to remain in the market. Such high-cost forecasters would earn smaller profits, but are not induced to exit if profits are nonnegative.

This potential for cost inefficiency is measured by the total profits which cost-efficient forecasters can earn under the incentive scheme.[20] This potential is not necessarily actualized, however. The actual extent of cost inefficiency from this source will depend on such subtleties as the distribution of forecaster cost types and the behavioral parameters which govern entry and exit into the markets for various forecasting tasks. Such a discussion is beyond the scope of this current paper.

It is likely that neither source of inefficiency can be readily eliminated by an unsophisticated forecast requisitioner. Eliminating the inefficient choice of the number of forecasters would require being able to compute VMP based on a comparison with the forecast that would be generated if the hiring of one less forecaster caused the remaining forecasters to exert greater efforts. If the requisitioner were sufficiently sophisticated, this problem of excessive entry might be remedied by making an adequate downward adjustment in forecaster compensation. The other problem, deterring the entry of high-cost forecasters, might be solved by reducing the excess profits of forecasters by means of a bidding scheme for the rights to submit forecasts to the forecast requisitioner. Such a bidding scheme is not straightforward, if forecasters differ in their optimal $\tau$'s, but may be feasible for a sophisticated requisitioner (perhaps with some distortion of incentives).[21] Such schemes will not be discussed further here, in our emphasis on what is possible for an unsophisticated principal.

## VII. Conclusions

Payment according to a very close proxy of VMP ensures that each forecaster will exert optimal levels of effort and causes a nearly optimal number of forecasters to be attracted to the forecasting task. In a competitive market with thousands of forecasters, it is expected that many forecasters will join intermediating firms and partnerships as a way of obtaining capital and insurance. By conditioning forecaster compensation on additional information, these intermediaries can furnish effectively risk-neutral forecasts to the forecast requisitioner, even when the risk-averse forecasters might otherwise furnish biased forecasts.

APPENDIX A. PROOF OF PROPOSITION 3.

To determine whether this can be a Nash equilibrium, suppose all other forecasters submit $T_i=\tau_i$ and $X_i(T_c)=ß_c I_i$ where $ß_c=\sigma_s^2/(\sigma_s^2+1/T_c)$. We then ask whether it is optimal for a particular forecaster to abide by the same strategy. Define:

$$I_c = \sum_{i=1}^{N} \tau_i I_i \;/\; \tau_c$$

$$\tau_{ci} = \tau_c - \tau_i$$

$$I_{ci} = \sum_{j\#i} \tau_j I_j \;/\; \tau_{ci} \qquad \text{(A.1)}$$

$$E_c = I_c - S$$

$$E_{ci} = I_{ci} - S$$

If we break down the variables in (21) into their component parts we obtain:

$$P_i(T_i,X_i,...) \qquad \text{(A.2)}$$

$$= 2(S+E_a)\{X_iT_i/(T_i+\tau_{ci})+ß_c(S+E_{ci})\tau_{ci}/(T_i+\tau_{ci})-ß_{ci}(S+E_{ci})\}$$

$$+ ß_{ci}^2(S+E_{ci})^2 - \{X_iT_i/(T_i+\tau_{ci})+ß_c(S+E_{ci})\tau_{ci}/(T_i+\tau_{ci})\}^2$$

where $ß_c = \sigma_s^2/(\sigma_s^2+1/(T_i+\tau_{ci}))$

and $ß_{ci} = \sigma_s^2/(\sigma_s^2+1/\tau_{ci})$

Taking expectations we obtain:

$$E(P_i) = 2ß_iI_i\{X_iT_i/(T_i+\tau_{ci})+ß_cß_iI_i\tau_{ci}/(T_i+\tau_{ci})-ß_{ci}ß_iI_i\} \qquad \text{(A.3)}$$

$$+ 2(ß_i/\tau_i)\{ß_c\tau_{ci}/(T_i+\tau_{ci})-ß_{ci}\} + ß_{ci}^2\{ß_i^2I_i^2+ß_i/\tau_i+1/\tau_{ci}\}$$

$$- X_i^2T_i^2/(T_i+\tau_{ci})^2 - 2X_iß_cß_iI_iT_i\tau_{ci}/(T_i+\tau_{ci})^2$$

$$- ß_c^2\{ß_i^2I_i^2+ß_i/\tau_i+1/\tau_{ci}\}\tau_{ci}^2/(T_i+\tau_{ci})^2$$

where $ß_i = \sigma_s^2/(\sigma_s^2+1/\tau_i)$

First-order conditions for maximization of expected pay require:

$$(\partial P/\partial X_i) = 2ß_iI_iT_i/(T_i+\tau_{ci}) - 2X_iT_i^2/(T_i+\tau_{ci})^2 \qquad \text{(A.4)}$$
$$- 2ß_cß_iI_iT_i\tau_{ci}/(T_i+\tau_{ci})^2 = 0$$

$$(\partial P/\partial T_i) = 2ß_iI_iX_i/(T_i+\tau_{ci}) - 2ß_iI_iX_iT_i/(T_i+\tau_{ci})^2 \qquad \text{(A.5)}$$
$$- 2ß_i^2I_i^2ß_c\tau_{ci}/(T_i+\tau_{ci})^2 - 2ß_i\tau_{ci}ß_c/(\tau_i(T_i+\tau_{ci})^2)$$
$$- 2T_iX_i^2/(T_i+\tau_{ci})^2 + 2T_i^2X_i^2/(T_i+\tau_{ci})^3$$
$$- 2\tau_{ci}X_iß_cß_iI_i/(T_i+\tau_{ci})^2 + 4T_i\tau_{ci}X_iß_cß_iI_i/(T_i+\tau_{ci})^3$$
$$+ 2\tau_{ci}^2ß_c^2(ß_i^2I_i^2+ß_i/\tau_i+1/\tau_{ci})/(T_i+\tau_{ci})^3$$
$$+ 2ß_i^2I_i^2\tau_{ci}ß_c/\{(T_i+\tau_{ci})^2[(T_i+\tau_{ci})\sigma_s^2+1]\}$$
$$+ 2ß_i\tau_{ci}ß_c/\{\tau_i(T_i+\tau_{ci})^2[T_i+\tau_{ci})\sigma_s^2+1]\}$$
$$- 2T_i\tau_{ci}X_iß_cß_iI_i/\{T_i+\tau_{ci})^3[(T_i+\tau_{ci})\sigma_s^2+1]\}$$
$$- 2\tau_{ci}^2ß_c^2[ß_i^2I_i^2+ß_i/\tau_i+1/\tau_i]/\{(T_i+\tau_{ci})^3[(T_i+\tau_{ci})\sigma_s^2+1]\}$$
$$= 0$$

It can be verified by substitution that $T_i=\tau_i$ and $X_i=ß_cI_i$ solves (A.4) and (A.5). Hence, it is a Nash equilibrium for all forecasters to submit $T_i=\tau_i$ and $X_i=ßI_i$, which is socially ideal.

## FOOTNOTES

1. NOTICE OF PATENT ISSUED: This paper describes a method of economic incentives, involving plural-forecaster payment systems, upon which the author and inventor has been issued a patent. (U.S. Patent 5,608,620) The patent on this invention only restricts actual use of the described invention; it does not restrict in any way the verbal or written discussion, description, or criticism of that invention.

2. See Osband (1985) Section 5.6 for a discussion of this restriction.

3. Forecasts of corporate profits and environmental costs would normally require the use of proxies to judge the success of current predictions. See Lundgren (1994) for an application of the VMP method to predict or estimate the value of an unobservable variable.

4. Let $I_i$ be the information set available to forecaster i and let $I_c$ be the combined information set of all n forecasters (but not necessarily known by any individual forecaster). Then $f(X_a)$ is conditional on $I_c$: $f(X_a)=f(X_a|I_c)$. $f(X_a)$ is not conditional on $X_c$ or $G(X_c)$, since the collective forecast might or might not reflect fully and accurately the information sets available to forecasters, depending on the incentives faced by forecasters.

5. Samuelson (1957, p. 209) suggests that financial speculators are generally not rewarded according to VMP. The incentive scheme suggested here attempts to remedy this deficiency.

6. The marginal contribution might or might not be equivalent to VMP, depending on how VMP is defined. The measurement of VMP requires a calculation of the net benefits which exist with forecaster i's predictions minus the net benefits which would exist in the absence of forecaster i's prediction. If we assume no replacement of forecaster i and no change in the effort levels of other forecasters as a result of the absence of forecaster i's prediction, then $MC_i=VMP_i$. In the absence of knowledge of how these other factors would affect the value of benefits in the absence of forecaster i's prediction, $MC_i$ is probably a good proxy for $VMP_i$.

7. The information set, Ii, refers to all bases for rational forecasting, including both objective data and subjective judgement. Certainly, different humans can interpret the same data quite differently--forecasting is not a pure mechanical or mathematical process.

8. We assume that each $X_i$ is a sufficient statistic for its corresponding $I_i$. If each $X_i$ is a single-valued prediction, this

condition might not hold. Under such circumstances, the forecast requisitioner may either request a vector of reports which are sufficient statistics, or else make do with an inferior (but less complicated) aggregator function.

9. An optimal prediction aggregator function is not necessarily unique. For example, if a weighted average of predictions is optimal, then the same weighted average times two is likewise optimal, provided that forecasters would simply submit the same predictions divided by two.

10. In a more generalized formulation of the utility function, we would also wish to take into account forecaster efforts, time, and money costs. If forecasters are paid on the margin according to VMP (k=1 in equation 6), then these other costs and efforts will be optimally determined by the forecaster. (See Section V.)

11. This assumption is reasonable, but is not critical to the proof. Theoretically, the forecast requisitioner could (perversely) choose a prediction aggregator function where this condition does not hold. For example, G(H,H,...,H)=2H. Forecasters would then be motivated to choose H=G*/2, so that when all predictions are doubled by the requisitioner, the result is G*.

12. The second-order condition for utility-maximization is also satisfied, provided forecasters are risk averse or risk neutral. Forecasters may also be risk seeking, providing they are not too risk seeking.

13. If $I_{ci}$ is a vector, there will be several integrations to correspond to each of the information variables of each rival forecaster. f( ) and g( ) are probability density functions conjectured by forecaster i and conditional on the information sets indicated. $g(I_{ci}|I_i)$ is the conjectured distribution of other forecasters' information, given forecaster i's observation of his own information. $f(X_a|I_i,I_{ci})$ is forecaster i's conjecture of the distribution of $X_a$, given his own information and the conjectured information of others. No presumption is made that any forecaster knows the information set of any other forecaster.

14. If the forecasters are risk neutral, the second-order conditions are assured. If forecasters are risk averse, the second-order conditions are also assured at this point, even though the first-order conditions might not be.

15. We may render a risk-averse forecaster effectively risk neutral if, instead of paying P, we pay V(P), where V(P) is chosen such that U(V(P)) is proportional to P. This requires ascertaining the degree of risk aversion in the forecaster's utility function. One possible disadvantage is that a weakly

risk-averse forecaster might be rendered risk seeking, if the actual degree of risk aversion is unknown.

16. It is still necessary to pay forecasters in accordance with the biased collective prediction, but any other use can be debiased, if the biasing factor is known or estimable. One way of estimating this bias might be to ask forecasters (independently of compensation) to furnish risk-neutral predictions, in addition to their risk-averse predictions.

17. The optimal combination of information is based on standard statistical theory. The derivation is not shown here.

18. The pay schedule is equivalent to assuming $F=0$ and $k=1$ in equation (6).

19. In a simulation of 2,178 cases, the mean value of the inefficiency due to attracting an excess number of forecasters was approximately 0.32% of forecaster costs. Approximately 55% of the sampled cases yielded a situation with $N_e=N_o$, while only 45% of the cases yielded $N_e>N_o$.

20. The mean value of potential inefficiency due to the excess profits arising from integer constraints was approximately 3.2% of forecaster costs in the previously mentioned sample of 2,178 cases.

21. Eliminating excess profits requires setting $F<0$ in (6). Attempting to do this with a bidding scheme (with bids proportional to $T_i$) may cause the effective k in (6) to vary away from 1, thereby inducing incorrect effort levels.

## REFERENCES


Kadane, Joseph B. and Robert L. Winkler (1988). "Separating Probability Elicitations From Utilities," Journal of the American Statistical Association, June, vol. 83, no. 402, pp. 357-363.

Lundgren, Carl (1994). "Peer-Group Forecasting Incentives for Unobserved Variables," manuscript, December.

Osband, Kent Harold (1985). Providing Incentives for Better Cost Forecasting, Ph.D. Dissertation, University of California, Berkeley. (Ann Arbor, MI: United Microfilms International)

Osband, Kent (1989). "Optimal Forecasting Incentives," Journal of Political Economy, vol. 97, no. 5, pp. 1091-1112.

Page, Talbot (1988). "Pivot Mechanisms as a Link between Probability and Preference Revelation," Journal of Economic Theory, vol. 44, pp. 43-62.

Samuelson, Paul A. (1957). "Intertemporal Price Equilibrium: A Prologue to the Theory of Speculation," Weltwirtschaftliches Archiv, vol. 79, no. 2, pp. 181-221.