Edinburgh School of Economics

Discussion Paper Series

Number 24

A Model of Management Teams

Eduardo Ley (Resources for Future, Washington DC)
Mark F J Steel (University of Tilburg)

Date
July 1995

Published by
School of Economics
University of Edinburgh
30 -31 Buccleuch Place
Edinburgh EH8 9JT
+44 (0)131 650 8361
http://www.ed.ac.uk/schools-departments/economics

THE UNIVERSITY *of* EDINBURGH

A Model of Management Teams

Eduardo Ley
Resources for the Future, Washington DC, USA
Mark F.J. Steel
Tilburg University, The Netherlands

April 1993
Version: July 19, 1995

Abstract. We present a simple model of management teams where the time it takes to make decisions is related to the size of the committee. We characterize the situations where larger or smaller sizes of the management team are desirable depending on the covariance structure of the signals that the managers observe.

Keywords. Management Team Size, Committee Size, Kalman Filter

JEL Classification System. D71, D89, L20

Address. Eduardo Ley: Resources for the Future, 1616 P St NW, Washington DC 20036. Email: ley@rff.org. Mark Steel: CentER and Department of Econometrics, Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands. Email: steel@kub.nl.

"We'd get the people from research, sales and production together, and everyone would say 'Not this' and 'Not that.' We'd talk but there would be no agreement. (...) Product planning has to be on a tight schedule. But we'd have another discussion, and another study and then more preparation. And finally, the decision would come months later." [Nobuhiko Kawamoto, President of Honda Motor Co. (*The Wall Street Journal*, 4/11/91, page A1.)]

"Despite what textbooks say, most important decisions in corporate life are made by individuals, not by committees. (...) To sum up: nothing stands still in this world. I like to go duck hunting, where constant movement and change are facts of life. You can aim at a duck and get it in your sights, but the duck is always moving. *In order to hit the duck, you have to move your gun.* But a committee faced with a major decision can't always move as quickly as the events it's trying to respond to. By the time the committee is ready to shoot, the duck has flown away." [Lee Iacocca, former head of Ford Motor Co. and of Chrysler Motor Co. (Lee Iacocca (1984): *Iacocca: An Autobiography*, New York: Bantam; page 52.)]

1. Introduction

Suppose that the profits a firm makes are positively related to *how well* a firm predicts the future — *i.e.*, other firms' choices, macroeconomic conditions, consumers' attitudes. We envision the firm as functioning with some default rules, strategies or policies until some changes are decided by the management team. The management makes evaluations of the state of nature, and takes decisions as to whether to alter its policy — *i.e.*, the *status quo*. We assume that changes in policy are given by some fixed reaction function that maps the pair consisting of the current evaluation of the state of nature and the present policy into some specific actions. Until another change in policy occurs, then the firm is in "automatic-pilot mode." A better evaluation of the state of nature is assumed to be associated with higher profits.

The state of nature, θ, is assumed to be constantly evolving as described by the following first-order Markov process

$$\theta_t = G_t\theta_{t-1} + w_t \tag{1}$$

where G_t is a known quantity and w_t is the system equation error. Each member, i, of the management team, or committee, can observe a signal $y_{i\tau}$ (which may

We thank Dallas Burtraw and an anonymous referee for useful comments.

be a vector or a scalar) which is linearly related to the unobservable vector θ_τ that represents the true state of nature at time τ,

$$y_{i\tau} = f'_{i\tau}\theta_\tau + v_{i\tau} \tag{2}$$

where $f_{i\tau}$ is known, and $v_{i\tau}$ is manager i's observation error.

Equations (1) and (2) can be embedded into a Kalman filter model. The Kalman filter model (Kalman (1960)) is a recursive procedure for making inferences about a vector of parameters using the history of some observable quantities which are linearly related to it. The density of the state parameters (the state of nature, θ) is evaluated at the current time, future observations are predicted, and, when they become available, the prediction error is used to update our inference about the state parameter. Under normality assumptions, the Kalman filter features some optimal statistical properties, and its simplicity and elegance makes it a very attractive algorithm. (See Meinhold and Singpurwalla (1983) for an introductory exposition.)

In the Kalman filter literature, we normally have $\tau = t$, and equations (1) and (2) are known as the *system equation* and the *observation equation*, respectively. In the usual Kalman filter formulation, there is an observation of y_{it} at each t and the posterior distribution of θ given the data is updated whenever a new observation becomes available — *i.e.*, at each t.

Suppose that when a team consists of n people, decisions cannot be taken at every t and the *updating step* is performed only every $\delta(n)$ periods. This means that equation (2) holds only for $\tau = \delta(n), 2\delta(n), \ldots$. On the other hand, the precision of the posterior distribution of θ increases with the team size, n. There is a tradeoff then between getting *more* information *less often* versus getting *less* information *more often*. We shall assume that $\delta(1) = 1$ and $\delta(n)$, which we call *delay function*, increases with n. In addition, for simplicity, we shall also assume that $\delta(n)$ is an integer for all n.

The motivation is that group decision-making is a long, expensive and complicated process and, while the management team is engaged in this task, the firm is in automatic-pilot mode. A single-manager team might generate less precise estimates but might be able to produce them more often.

In statistical terms, the question becomes whether it is better to get a bigger sample each time we sample (not at every t) and update less often, or to sample less at each t but update more frequently. Put differently; which is the optimal group size, n^*, that minimizes posterior uncertainty?

In this paper we shall characterize different aspects of the covariance matrix of the signal errors that the committee members receive which are important to the optimal size of the management team.

In the simple model that we develop in this paper, we don't address any problems related to the optimal decision rule by a committee under different environments. This is an important issue and its study requires imposing some structure on the type of problems that the committee is facing (see, *e.g.*, Koh (1994), and Sah and Stiglitz (1988)).[1] We also ignore any communication and information-pooling problems (see, *e.g.*, Berger (1985), Section 4.11 on 'Combining Evidence and Group Decisions' and the references cited therein). In the model developed in this paper, information from different sources is pooled via the likelihood by the usual Bayesian updating mechanism.

2. The Management Model

For simplicity, we will assume that each manager gets a single signal — *i.e.*, that $y_{i\tau}$ is a scalar and not a vector — but our results extend to the general case where each manager observes a vector of signals. The system and observation equations are:

$$\theta_t = G_t \theta_{t-1} + w_t, \qquad t = 1, 2, \ldots \tag{3}$$

$$y_\tau = F'_\tau \theta_\tau + v_\tau, \qquad \tau = \delta(n), 2\delta(n), \ldots \tag{4}$$

The number of managers, n, is reflected in the dimension of y_τ and in the frequency of the observation equation (4) through the delay function, $\delta(n)$. More explicitly, we have $F_\tau = [f_{1\tau}, f_{2\tau}, \ldots, f_{n\tau}]$, and $y_\tau = [y_{1\tau}, y_{2\tau}, \ldots, y_{n\tau}]'$.

The distributional assumptions are[2]

$$w_t \sim N(0, W),$$
$$v_t \sim N(0, V).$$

[1] See also Huberman and Loch (1994) for a derivation of bounds on team size based on the members' incentives to share information; and Katzner (1995) for a model in which the cost of making and implementing decisions depend on the fraction of employees who participate in the process and in the similarity of their preferences.

[2] Actually, we don't need to require normality. Girón and Rojano (1994) show that ellipticity is enough for the Kalman filter to have the usual recursion equations. In that case, a diagonal covariance matrix implies lack of correlation but not independence. Covariance matrices that are not constant over time can be easily accommodated.

The prior distribution on θ, corresponding to our knowledge at time 0, before any observation is taken, is

$$\theta_0 \sim N(\mu_0, C_0).$$

The information in the observations acquired through (4) and the dynamics of θ inherent in (3) will jointly lead to the posterior, given a set of observations Y_t,

$$\theta_t | Y_t \sim N(\mu_t, C_t). \tag{5}$$

Thus, at time t all available information concerning the state of nature θ_t is summarized by its conditional mean and variance.

2.1. Single Manager

If $n = 1$ (which implies $\delta(1) = 1$ by assumption), then we use

$$\begin{aligned} e_t &= y_t - F_t' G_t \mu_{t-1} \\ R_t &= G_t C_{t-1} G_t' + W \\ Q_t &= F_t' R_t F_t + V \end{aligned}$$

to update the posterior density of θ_t given by (5),

$$\begin{aligned} \mu_t &= G_t \mu_{t-1} + R_t F_t Q_t^{-1} e_t \\ C_t^{-1} &= R_t^{-1} + F_t V^{-1} F_t'. \end{aligned}$$

These are the usual Kalman recursion equations —see, *e.g.*, Meinhold and Singpurwalla (1983) or Girón and Rojano (1994).

2.2. Management Team

When $\delta(n) \neq 1$, but some other integer value, we only observe the y's at $\tau = \delta(n), 2\delta(n), \ldots, l\delta(n)$. Then

$$\begin{aligned} e_\tau &= y_\tau - F_\tau' G_\tau \mu_{\tau-1}, \\ R_t &= G_t C_{t-1} G_t' + W, \\ Q_\tau &= F_\tau' R_\tau F_\tau + V. \end{aligned}$$

Note that the updating of R_t is performed through the system equation and thus takes place every period. In order to change the values of e_τ and Q_τ we require an observation which occurs every $\delta(n)$ periods. Before a new observation is made, we still use (3) at each t, and after a new observation we update the posterior density of θ_t. The posterior mean is given by

$$\mu_t = \begin{cases} G_t\mu_{t-1}, & \text{for } t \neq \tau \\ G_\tau\mu_{\tau-1} + R_\tau F_\tau Q_\tau^{-1} e_\tau, & \text{for } t = \tau; \end{cases}$$

and the precision is updated by

$$C_t^{-1} = \begin{cases} R_t^{-1}, & \text{for } t \neq \tau \\ R_\tau^{-1} + F_\tau V^{-1} F_\tau', & \text{for } t = \tau. \end{cases} \tag{6}$$

Given $\delta(n)$, now how to choose n such that the posterior precision, C_t^{-1}, is maximum? In a classical context, this corresponds to the n which minimizes MSE, if we use μ_t as a point estimate for θ_t.

3. Covariance Structures for the Managers' Signals

If all managers have the same *observation models*, then

$$\begin{aligned} F_\tau &= [f_{1\tau}, \ldots, f_{n\tau}] = [f_\tau, \ldots, f_\tau] \\ &= f_\tau \otimes \iota_n' = f_\tau \iota_n' \end{aligned} \tag{7}$$

where ι_n is a column vector of n ones.[3]

In a sense, R_τ^{-1} in (6) is the precision before incorporating the last information available, and the value of the last observation is

$$S_\tau \equiv F_\tau V^{-1} F_\tau' = (f_\tau \otimes \iota_n') V^{-1} (f_\tau' \otimes \iota_n).$$

Let us now consider possible specifications for V, the covariance matrix of the observation errors of the team.

It is worth noting that, by assumption, the delay function, $\delta(n)$, is a function of the size of the team only. It does not depend on the similarity of the views

[3] The operator '$\otimes$' is the *Kronecker product* which involves multiplying each element of the matrix of the left by the entire matrix of the right.

of the committee members at all. In the sampling analogy, the cost of sampling would not be affected by the correlation among samples. It is conceivable that, in real life, meetings would be shorter, the greater the similarity in the views of the members of the committee. However, $\delta(n)$ is not intended to represent only the length of the actual meetings that take place but all the costs incidental to having large management teams.

3.1. Independent Errors

A simple hypothesis is that every manager commits errors independently of those of the other members of the team. In addition, the quality of all managers could be considered equal in that all errors have equal variances. In this case,

$$V = \sigma^2 I_n$$

where I_n is an identity matrix of dimension n, and

$$S_\tau = \sigma^{-2}(f_\tau f_\tau' \otimes \iota_n' \iota_n) = n\sigma^{-2} f_\tau f_\tau'.$$

3.2. Equicorrelated Errors

If we maintain the assumption that all managers have equal observation variance, but allow for the same amount of correlation between all members of the team, we get the second case of interest. If

$$V = (\sigma^2 - \gamma)I_n + \gamma \iota_n \iota_n' \tag{8}$$

with $\sigma^2 > \gamma$ for $\gamma \geq 0$ and $\sigma^2 > -(n-1)\gamma$ for $\gamma < 0$, we obtain

$$V^{-1} = (\sigma^2 - \gamma)^{-1} I_n - \frac{\gamma}{(\sigma^2 - \gamma)(\sigma^2 - \gamma + n\gamma)} \iota_n \iota_n'$$

and the value of the last observation becomes

$$\begin{aligned} S_\tau = (f_\tau \otimes \iota_n')V^{-1}(f_\tau' \otimes \iota_n) &= \frac{n}{\sigma^2 + (n-1)\gamma} f_\tau f_\tau' \\ &< \frac{n}{\sigma^2} f_\tau f_\tau' \qquad \text{if } \gamma > 0 \text{ and } n > 1. \end{aligned}$$

Of course, this case specializes to independence for $\gamma = 0$. Note that the correlation coefficient between team members is given by $\rho = \gamma / sigma^2$ and that we need to restrict ρ to the interval $(-(n-1)^{-1}, 1)$ in order to assure that V in (8) is positive definite.

4. Firm Policy

In this model, all teams —regardless of size— optimally process the information available to them; viewed from a classical perspective, the models produce efficient unbiased forecasts for θ. Therefore, a *better evaluation* of the state of nature is synonymous with a higher precision, C^{-1}. Then, when comparing the performance of a single manager against a management team we only need to compare the precision of their posterior densities of the state of nature θ at every t. This is what we call *overall firm policy.* We might also want to consider the case where decisions are only taken every $\delta(n)$ periods, even if the management consists of a single manager who updates her view of the world with a higher frequency. We call this case *special decisions.*

Table 1. Comparison of precision matrices at each t: one manager vs n-sized team.

t	R_t	C_t^{-1}	
		$n = 1$	$n > 1$
1	$G_1 C_0 G_1' + W$	$R_1^{-1} + ff'\sigma^{-2}$	R_1^{-1}
2	$G_2 C_1 G_2' + W$	$R_2^{-1} + ff'\sigma^{-2}$	R_2^{-1}
...	...	...	...
$\delta(n)$	$G_\delta C_{\delta-1} G_\delta' + W$	$R_\delta^{-1} + ff'\sigma^{-2}$	$R_\delta^{-1} + nff'(\sigma^2 + (n-1)\gamma)^{-1}$

Table 2. Comparison of precision matrices at each t: n vs $n + 1$.

t	R_t	C_t^{-1}	
		n	$n + 1$
1	$G_1 C_0 G_1' + W$	R_1^{-1}	R_1^{-1}
2	$G_2 C_1 G_2' + W$	R_2^{-1}	R_2^{-1}
...	...	...	...
$\delta(n)$	$G_{\delta(n)} C_{\delta(n)-1} G_{\delta(n)}' + W$	$R_{\delta(n)}^{-1} + nff'(\sigma^2 + (n-1)\gamma)^{-1}$	$R_{\delta(n)}^{-1}$
...	...	...	...
$\delta(n+1)$	$G_{\delta(n+1)} C_{\delta(n+1)-1} G_{\delta(n+1)}' + W$	$R_{\delta(n+1)}^{-1}$	$R_{\delta(n+1)}^{-1} + (n+1)ff'(\sigma^2 + n\gamma)^{-1}$

In tables 1 and 2, the case of independence corresponds to $\gamma = 0$. Note that the series of C_t^{-1} matrices generated by the last two columns differ in that the C_t^{-1} used in the formula for C_t either incorporates the information gain of an observation ($n = 1$) or not ($n > 1$ and $t < \delta(n)$).

In case $G_t C_{t-1} G_t' + W$ would stay roughly the same for both choices of n, however, we can make some simple comparisons. This applies, in particular, if $ff'\sigma^{-2}$, the added precision of one single manager observation, is small compared to $(G_tC_{t-1}G_t' + W)^{-1}$ which represents the accumulated information of the firm at time t without this last observation. Of course, this situation would occur if R remained constant, but it is only required that R evolves similarly under both single managers and team decisions. We are then able to use tables 1 and 2 for deriving some "rules of thumb."

4.1. Overall Firm Policy

Here we compare the sum of the precision increases over all periods until a certain horizon $\tau = \delta(n)$ (or a multiple of $\delta(n)$). Of course, different values of n will also lead to different updates of R at time t. The value of R can go up, or down with updating; if R would not be affected greatly by the value of n,[4] then we could compare directly the cases with $n = 1$ and $n > 1$ over, *e.g.*, the period from $t = 1$ to $t = \delta(n)$.

In that case, then the n-manager team would be preferable to a single manager whenever

$$\frac{\delta(n)(\delta(n)+1)}{2\sigma^2} < \frac{n}{\sigma^2 + (n-1)\gamma}$$

where $\gamma = 0$ reflects independence; or, rearranging, if

$$\frac{\delta(n)(\delta(n)+1)}{2} < \frac{n}{1 + \frac{(n-1)\gamma}{\sigma^2}} = \frac{n}{1+(n-1)\rho} \equiv a(n), \tag{9}$$

where $\rho = \gamma/\sigma^2$ which measures the correlation between the managers' signals. Note that $a(n)$ always increases with n, since we must have $-(n-1)^{-1} < \rho < 1$.

We can use table 2 to compare an n-sized team with an $(n+1)$-sized team. For fair comparison, consider the horizon $\delta(n) \times \delta(n+1)$, so that a size n team has made $\delta(n+1)$ decisions, and the team of size $n+1$ has had $\delta(n)$ updates. Again, if R evolves similarly, we would prefer to augment the size of the team from n to $n+1$ whenever

$$\frac{a(n)}{\delta(n)(\delta(n)+1)} < \frac{a(n+1)}{\delta(n+1)(\delta(n+1)+1)}.$$

[4] For R to remain constant, consider the case where both G and C remain constant (from the second column of table 1). For $n > 1$ that implies that $R = GC(n)G' + W$. For $n = 1$ we must have $C(1)^{-1} = (GC(1)G' + W)^{-1} + ff'\sigma^{-2}$. We write $C(n)$ to stress that the precision matrices at each t will differ in the $n = 1$ and $n > 1$ cases.

An analogous formula applies for comparing sizes n and m, and for $m = 1$ this will reduce to (9), since $a(1) = \delta(1) = 1$.

4.2. Special Decisions

In the case of crucial decisions that need to be taken at a particular point in time τ, we can compare the information accumulated at τ by updating every period with that obtained by a team making only one decision, namely at time τ. As the team is specifically put together for this decision, we take $\tau = \delta(n)$. In the equicorrelated case, when $n = 1$ after $\delta(n)$ updates,

$$C_\tau^{-1} = \delta(n)\sigma^{-2} f f' + R^{-1}$$

and when $n > 1$, after one update

$$C_\tau^{-1} = \frac{n}{\sigma^2 + (n-1)\gamma} f f' + R^{-1},$$

so that $n > 1$ is preferable to $n = 1$ if

$$\delta(n) < a(n). \tag{10}$$

If we now compare teams of sizes n and $n+1$, we should consider the precision at some time point corresponding to a crucial decision, say, $\tau = \delta(n) \times \delta(n+1)$. Then, we would favor size $n+1$ over size n if

$$\frac{a(n)}{\delta(n)} < \frac{a(n+1)}{\delta(n+1)},$$

and, generally, we prefer size n over size m if

$$\frac{a(m)}{\delta(m)} < \frac{a(n)}{\delta(n)}.$$

Choosing $m = 1$, we recuperate (10).

If C_{t-1} for $n = 1$ becomes larger than C_{t-1} with a team (*i.e.*, in case of updating every period, the precision decreases) then we will choose larger n than suggested by (9) or (10). In opposite cases we shall choose a smaller team.

4.3. Behavior of $a(n)$

We have seen that $a(n)$ as defined in (9) can, in some cases, be compared with $\delta(n)$ or $\delta(n)(\delta(n)+1)/2$ directly. Let us now examine the properties of $a(n)$ as n varies.

Clearly $a(1) = 1$ and for all possible values of ρ we have that $a(n)$ is strictly increasing in n. For positive ρ (*i.e.*, $0 < \rho < 1$) we obtain

$$\lim_{n\to\infty} a(n) = \frac{\sigma^2}{\gamma} = \frac{1}{\rho} > 1$$

whereas for other values of ρ (*i.e.*, $-(n-1)^{-1} < \rho \leq 0$),

$$\lim_{n\to 1+1/|\rho|} a(n) = \infty.$$

In this latter case, the requirement of a positive definite observation error matrix V induces an asymptote at $n = 1 + |\rho|^{-1}$. This can be interpreted as follows. Given σ^2 and a negative error correlation we cannot let the team grow too large, as then this fixed negative correlation would confuse matters such that an observation would have a negative value. If we want n to grow, we need to decrease $|\rho|$.

As managers tend to observing without error, ρ will tend to one, and $a(n)$ will tend to one for any value n is allowed to take. In that case, we will, of course, opt for a single manager (since $\delta(n) \geq 1$ if $n > 1$).

4.4. Simple Rules of Thumb

On the basis of (9) and (10), some rules of thumb can now be deduced, which are valid if the values of R_t for teams and single managers are the same (and approximately valid if they are close). Without making any assumptions on $\delta(n)$, except that $\delta(1) = 1$ and $\delta(n)$ increases with n and is bounded for all finite n, we can now state:

> **Rule 1.** *With* **positive** ρ $(0 < \rho < 1)$ *one should never choose* n *such that* $\delta(n) > \rho^{-1}$ *for a special decision or* $\delta(n)(\delta(n)+1)/2 > \rho^{-1}$ *for overall policy.*

> **Rule 2.** *With* **negative** ρ $(-(n-1)^{-1} < \rho < 0)$ *one should always choose* n *slightly smaller than* $1 + |\rho|^{-1}$, *provided such a value is feasible.*

The first rule tells us that, since the value of $a(n)$ can never attain its limit, ρ^{-1}, you should never choose a team so large that it takes ρ^{-1} or more time periods to reach a conclusion for a special decision. For overall firm policy you even choose a much smaller team in general. The more people's views overlap (the higher ρ) the smaller the optimal team will be.

The second rule of thumb derives from the fact that $a(n)$ increases without bounds as $n \to 1 + |\rho|^{-1}$ and thus can be made greater than the bounded $\delta(n)$ or $\delta(n)(\delta(n)+1)/2$ by as much as we want. As γ becomes more and more negative, managers' views diverge too much and we need to reduce the team size. However, a slight negative correlation can lead to the existence and optimality of relatively large teams. An effect of "complementarity" of managers can greatly increase the value of the team. Note that this second rule is not affected by the type of decisions that have to be taken. It always seems to pay to put together a team of managers that view issues from somewhat different angles. Remark that $1 + |\rho|^{-1}$ can be quite large for small negative correlations and teams of such size could be infeasible in practice. In cases where *Rule 2* indicates an unreasonably large value for n, we can resort to the formulas in Subsections 4.1 and 4.2 for further guidance.

5. Illustrative Example

In this section we do some simulations using the model in section 6.2 of Meinhold and Singpurwalla (1983). In particular, we use:

$$\theta_t = \frac{1}{2}(-1)^t \theta_{t-1} + w_t, \qquad w_t \sim N(0,1).$$

We set $f_\tau = 1$ in equation (7), and $\sigma = 2$ in equation (8) where we allow for different values of γ. We show simulations for $\gamma = -0.1$, -0.025, 0, and 3.2 which imply values for ρ of -0.025, -0.00625, 0 and 0.8.[5] We also use various specifications for $\delta(n)$. We choose $\mu_0 = \theta_0 = 1$, and $C_0 = I$. To make the problem manageable,[6] we set the maximum possible team size, $\bar{n}$, equal to 15. The simulations illustrate a wide variety of cases. While our approximate

[5] We remind the reader that, with $\bar{n} = 15$, ρ must be between $-14^{-1} = -0.071$ and 1.

[6] Note, for example, that with n managers we must generate a Normal vector of dimension n with covariance V. In addition, the possible range of negative values for γ decreases as n increases.

formulas and rules of thumb work well in most cases, we also show situations were some of the rules could possibly mislead us.

As for any team size, the procedure makes optimal use of the information on the state of the world, θ, we shall focus here on the precision of the team's forecast. In particular, since we are going to focus on firm policy,[7] we average the posterior precision of a size-n team forecasts from $3\delta(n)+1$ to $4\delta(n)$ (to get away from the filter initialization). The results are shown in tables 3 and 4.[8] Table 3 shows the optimal team size for different values of $\delta(n)$ and ρ for the overall firm policy.

Let us now compare these actual values of n^* resulting from running the Kalman filter with the values suggested by the approximate formulas in Subsections 4.1 and 4.2. Comparing teams of size n with those of $n+1$ members, the approximate rule in Subsection 4.1 successfully locates all the local optima for $\rho = -0.00625$, 0, and 0.8. Using the same rule to contrast size-n with size-m teams, we differ from the globally optimal n^* in table 3 for $\delta(n) = \lfloor \frac{3}{4}n \rfloor + 1$ with $\rho = -0.00625$ and $\rho = 0$. Thus, the approximation underlying this rule seems reasonably accurate in this simulation. However, for $\rho = -0.025$ the approximate rule would lead us astray.

Table 3. Optimal team size for overall firm policy.

$\delta(n)$	ρ			
	-0.025	-0.00625	0	0.8
n	$n^* = \bar{n}$	$n^* = 1$	$n^* = 1$	$n^* = 1$
$\lfloor \frac{3}{4}n \rfloor + 1$	$n^* = \bar{n}$	$n^* = 13$	$n^* = 5$	$n^* = 1$
$\lfloor \sqrt{n} \rfloor$	$n^* = \bar{n}$	$n^* = \bar{n}$	$n^* = \bar{n}$	$n^* = 3$

If we consider the evidence in table 3 in the light of the even simpler rules of thumb in Subsection 4.4, we note that for $\rho = 0.8$ *Rule 1* excludes $n > 1$ for $\delta(n) = n$ and $\delta(n) = \lfloor \frac{3}{4}n \rfloor + 1$, so that we would take exactly the optimal $n^* = 1$. For $\delta(n) = \lfloor \sqrt{n} \rfloor$ we find $n \leq 3$ according to *Rule 1*, which again corresponds to the actual n^*. For $\rho = -0.00625$, *Rule 2* would suggest $n = 160$, but this is

[7] For special decisions, in this example, it is always optimal to have a team of the maximum size which allows to update the filter prior to the time when the special decision has to be taken.

[8] Here $\lfloor x \rfloor$ denotes the largest integer no greater than x.

not a feasible team size (the maximum size is 15). Thus, *Rule 2* can not really guide us here. However, for the case with $\rho = -0.025$ *Rule 2* indicates $n = 40$ and here n^* is equal to the maximum size of 15. The rule would thus, at least, point us in the right direction in this case.

6. Conclusion

Group decision making is common in many corporate organizations. Since there might be decreasing marginal productivity in gathering and digesting information by single individuals, 'parallel processing' by committees is often used. However, as noted, among others, by Koh (1994), information exchange in this context is also costly even ignoring strategic issues within a management team.

In this paper we present a simple model where we make this tradeoff explicit and characterize situations where larger or smaller management teams might be desirable depending on characteristics of the covariance between the signals that the managers observe. In particular, we find that (1) with positive correlation between the managers' signals, the larger the correlation, the smaller the optimal size will be; and (2) with negative correlation, a slight negative correlation might lead to the existence of large management teams due to a complementarity effect.

An example illustrates the relative accuracy of simple approximate formulas for choosing team size, and also shows that even simpler rules of thumb may work well in some cases. Both our approximate formulas and the rules of thumb seem most reliable for cases with positive correlation between managers' observation errors. We conjecture that this would be the prevalent situation in practice.

Table 4. Average Forecast Precision for different values of ρ and $\delta(n)$.

	$\rho = -0.025$						$\rho = 0$					
	$\delta(n) = n$		$\delta(n) = \lfloor \frac{3}{4}n \rfloor + 1$		$\delta(n) = \lfloor \sqrt{n} \rfloor$		$\delta(n) = n$		$\delta(n) = \lfloor \frac{3}{4}n \rfloor + 1$		$\delta(n) = \lfloor \sqrt{n} \rfloor$	
n	$\delta(n)$	$\bar{C}^{-1}$	$\delta(n)$	$\bar{C}^{-1}$	$\delta(n)$	$\bar{C}^{-1}$	$\delta(n)$	$\bar{C}^{-1}$	$\delta(n)$	$\bar{C}^{-1}$	$\delta(n)$	$\bar{C}^{-1}$
1	1	1.0590	1	1.0590	1	1.0590	1	*1.0590*	1	1.0590	1	1.0590
2	2	1.0599	2	1.0599	1	1.3572	2	1.0526	2	1.0526	1	1.3430
3	3	1.0604	3	1.0604	1	1.6585	3	1.0460	3	1.0460	1	1.6160
4	4	1.0621	4	1.0621	2	1.3708	4	1.0403	4	1.0403	2	1.3274
5	5	1.0652	4	1.1439	2	1.5338	5	1.0358	4	*1.1072*	2	1.4606
6	6	1.0695	5	1.1334	2	1.7038	6	1.0322	5	1.0886	2	1.5921
7	7	1.0750	6	1.1291	2	1.8821	7	1.0292	6	1.0757	2	1.7224
8	8	1.0814	7	1.1287	2	2.0699	8	1.0267	7	1.0663	2	1.8518
9	9	1.0887	7	1.1855	3	1.7651	9	1.0246	7	1.1031	3	1.5730
10	10	1.0970	8	1.1837	3	1.9055	10	1.0229	8	1.0911	3	1.6586
11	11	1.1061	9	1.1853	3	2.0547	11	1.0213	9	1.0816	3	1.7439
12	12	1.1162	10	1.1895	3	2.2138	12	1.0200	10	1.0740	3	1.8289
13	13	1.1273	10	1.2405	3	2.3838	13	1.0188	10	1.0994	3	1.9137
14	14	1.1394	11	1.2456	3	2.5662	14	1.0177	11	1.0907	3	1.9984
15	15	*1.1527*	12	*1.2534*	3	*2.7623*	15	1.0168	12	1.0835	3	*2.0829*
	$\rho = -0.00625$						$\rho = 0.8$					
	$\delta(n) = n$		$\delta(n) = \lfloor \frac{3}{4}n \rfloor + 1$		$\delta(n) = \lfloor \sqrt{n} \rfloor$		$\delta(n) = n$		$\delta(n) = \lfloor \frac{3}{4}n \rfloor + 1$		$\delta(n) = \lfloor \sqrt{n} \rfloor$	
n	$\delta(n)$	$\bar{C}^{-1}$	$\delta(n)$	$\bar{C}^{-1}$	$\delta(n)$	$\bar{C}^{-1}$	$\delta(n)$	$\bar{C}^{-1}$	$\delta(n)$	$\bar{C}^{-1}$	$\delta(n)$	$\bar{C}^{-1}$
1	1	*1.0590*	1	1.0590	1	1.0590	1	*1.0590*	1	*1.0590*	1	1.0590
2	2	1.0533	2	1.0544	1	1.3465	2	0.9237	2	0.9237	1	1.0914
3	3	1.0473	3	1.0494	1	1.6262	3	0.8705	3	0.8705	1	*1.1038*
4	4	1.0424	4	1.0455	2	1.3377	4	0.8422	4	0.8422	2	0.9334
5	5	1.0385	4	1.1157	2	1.4775	5	0.8246	4	0.8432	2	0.9355
6	6	1.0355	5	1.0988	2	1.6174	6	0.8126	5	0.8252	2	0.9369
7	7	1.0332	6	1.0876	2	1.7577	7	0.8040	6	0.8130	2	0.9379
8	8	1.0313	7	1.0798	2	1.8990	8	0.7974	7	0.8042	2	0.9387
9	9	1.0299	7	1.1205	3	1.6136	9	0.7923	7	0.8044	3	0.8767
10	10	1.0288	8	1.1102	3	1.7095	10	0.7882	8	0.7977	3	0.8770
11	11	1.0279	9	1.1024	3	1.8062	11	0.7848	9	0.7925	3	0.8773
12	12	1.0272	10	1.0965	3	1.9040	12	0.7819	10	0.7883	3	0.8775
13	13	1.0266	10	*1.1262*	3	2.0029	13	0.7795	10	0.7884	3	0.8777
14	14	1.0263	11	1.1193	3	2.1030	14	0.7774	11	0.7849	3	0.8779
15	15	1.0260	12	1.1138	3	*2.2042*	15	0.7756	12	0.7820	3	0.8780

References

Berger, J.O. (1985): *Statistical Decision Theory and Bayesian Analysis*. New York: Springer-Verlag.

Girón, F.J. and J.C. Rojano (1994): "Bayesian Kalman Filtering with Elliptically Contoured Distributed Errors," *Biometrika*, **81**, 390–395.

Huberman, B.A. and C.H. Loch (1994): "Collaboration, Motivation and the Size of Organizations," mimeo, INSEAD, Fontainebleau, France.

Kalman, R.E. (1960): "A New Approach to Linear Filtering and Prediction Problems," *Journal of Basic Engineering*, **82**, 34–45.

Katzner, D.W. (1995): "Participatory Decision-Making in the Firm," *Journal of Economic Behavior and Organization*, **26**:2, 221–236.

Koh, W.T.H. (1994): "Making Decisions in Committees: A Human Fallibility Approach," *Journal of Economic Behavior and Organization*, **23**:2, 195–214.

Meinhold, R.J. and N.D. Singpurwalla (1983): "Understanding the Kalman Filter," *The American Statistician*, **37**:2, 123–127.

Sah, R.K. and J.E. Stiglitz (1988): "Committees, Hierarchies and Polyarchies," *Economic Journal*, **98**, 451–470.