Hierarchies and The Self-Control of Risk Attitude

Luc Wathieu

Series No. MKTG 95.031
December, 1994

December 1994

HIERARCHIES AND THE SELF-CONTROL OF RISK ATTITUDE*

Luc Wathieu

*The Hong Kong University of Science and Technology
Clear Water Bay, Kowloon, Hong Kong*

* Thanks are due to Bernard Sinclair-Desgagné, Spyros Makridakis, Anil Gaba, Robin Hogarth, Bob Winkler, Saugata Banerjee, Jerry Ross, and seminar audiences at INSEAD, HKUST, FUR VII and EIASM. Financial support from ICM, the Wei Lun Fund and INSEAD is gratefully acknowledged.

Abstract

In hierarchies, top members are not continuously informed about the intermediary outcomes of their policies. This feature of organizational decision making is often regarded as a drawback caused by communication costs. On the contrary, we argue that such information decentralization creates a valuable interface between a decision maker and his environment. We assume a reference-dependent utility function à la Kahneman and Tversky. Personal contact with an intermediary outcome provokes a shift of reference point, and thereby a change of attitude towards the forthcoming alternative risks. Information decentralization stabilizes the reference point. We characterize classes of dynamic choice situations where this effect should prove desirable. Then, we show that hierarchical systems can help decision makers overcome their personal conservative tendencies.

Key words: information decentralization, reference points, dynamic choice under risk, conservatism

JEL classification numbers: L200, D800, D900

Chandler (1962) implied that, by decentralizing an organization, one can modify the psychological attitude of the involved decision makers:

"The basic reason for its [the M-form of organization] success was simply that it clearly *removed the executives* responsible for the destiny of the entire enterprise from the more routine operational activities, and so gave them the time, information and even *psychological commitment for long-term planning and appraisal*" (Chandler (1962), p.309, italic added).

The nature and implications of the relationship between decentralization and "psychological commitment" has received no attention from the decision making literature. However, these issues are crucial to understand both decision making in organizations and the organization of decision making. This paper makes a formal investigation of the link between decentralization and a decision maker's attitude in a dynamic choice context.

Basic Intuition

It is a well-documented fact that decision makers use reference points when they evaluate the possible outcomes of risky monetary alternatives. A reference point is a benchmark that determines whether an outcome is perceived as a loss or as a gain. Evaluation of losses and gains is typically asymmetric, since loss aversion is a frequent and dominant feature of human preferences (see, e.g., Tversky and Kahneman (1991)). The most common reference point, according to the literature, is accumulated wealth.

In a dynamic context, as wealth evolves, the reference point shifts and the utility associated with each wealth position changes. For instance, a position that initially corresponds to a gain can be evaluated as a loss after an increase in wealth has been perceived. The present paper examines the problem of dynamic choice under uncertainty for an expected utility maximizer whose reference point is subject to such volatility.

Suppose that you lose \$1000 on one of your personal investments. Loss aversion predicts that you experience a severe dissatisfaction. Shortly after, this investment yields a gain of \$1000. Now, when you consider your experienced utility over the period, you may find that you would have preferred to hear nothing about the intermediary outcome and be directly

exposed to a cumulated payoff of \$0. When such is the case, it means that you did not experience the \$1000 gain as a compensation for the earlier loss. This phenomenon corresponds to a shift of reference point. As this example illustrates, an individual with a volatile reference point should develop, ex ante, a preference relation over the frequency of feedback. Indeed, when such frequency is under control, it is possible to stabilize the reference point and the corresponding utility function for a while. Thus a meta-decision can be made, that determines future preferences among alternative risks.

Feedback, however, is not just stimulus for (dis)satisfaction and reference point shifts. It often conveys useful information about future alternatives. How, then, could one be able to process any such information while controlling exposure to intermediary outcomes? *The necessary device is hierarchical.* In hierarchies, decision makers choose policies, order their implementation and organize a calendar of appraisals. Subordinates are then in charge of processing information, executing the policy and reporting in due time. Because they allow temporary *information decentralization*,¹ hierarchies permit to differentiate between processed information and perceived information. Perceived information anticipatively motivates policy making at the top, whereas processed information permits contingent implementation at the lower level.

Thus our theory is that, by decentralizing information about the intermediary outcomes of his plans, a decision maker can (and often should) consistently achieve self-control of his risk attitude, while remaining capable of contingent planning. This paper is devoted to an analysis of the circumstances under which information decentralization and the implied stabilization of the reference point are desirable, in terms of the decision maker's expected utility. We will also examine some properties of the resulting choice behavior, and show that decentralization is a remedy against individual conservatism in a number of situations.²

¹ We borrow this term from Marschak and Radner (1972). When we write "(de)centralization", we always refer to "information (de)centralization".

² On conservatism in individual decision making, see Samuelson and Zeckhauser (1988).

Link with Organization Theory

That top members of hierarchies are not kept continuously informed about the intermediary outcomes of their policies is often regarded as an unavoidable drawback by organization theorists. For instance, Marschak and Radner (1972) relate the need for decentralization to communication costs. Here, in contrast, we argue that information decentralization can create a valuable interface between decision makers and their environment, which enables them to consistently adopt a desired risk attitude.

In that sense, we share the viewpoint originally expressed by Barnard (1968), according to which the "essence of formal organization" is to achieve "deliberate adoption of means and ends" that are out of reach for isolated individuals. However, we do not deal with the limitations of human capabilities (as Radner (1992) or Geanakoplos and Milgrom (1991) do, for instance). Rather, we focus on the peculiarities of individual risk attitudes that are motivated by what Tversky and Kahneman (1991) call "the psychophysics of hedonic experience."

If reference point shifts are considered as "biases", what we claim here is that hierarchies are debiasing systems (to be added to the list of Arkes (1991)), which decision makers should find attractive under some circumstances. Postrel and Rumelt (1992) have a similar philosophy, when they suggest that hierarchies offer a remedy to individual impulsiveness by imposing routines and habits on members.

Our emphasis on the transformation of risk attitude in hierarchies could be related to the observation by McCrimmon and Wehrung (1986) that decision makers are less risk averse in business context than in personal context.

Link with Dynamic Choice Theory

Our work can also be seen as an essay in dynamic choice under changing tastes. This field of research, that was initiated by Strotz (1956), primarily focused on addictive consumption goods such as tobacco, alcohol or food. Only recently, with a debate on the dynamic implications of non-expected utility theories, have we seen the appearance of a more general concern about time-inconsistent preferences (see, e.g., Machina (1989)). Strotz and his

followers have insisted on the structure of time preferences (i.e., non-exponential discounting) as the key explanatory variable for the changing taste phenomenon. The possible role of reference point shifts was only emphasized lately, by George Loewenstein (1988).

How to rationally decide under changing preferences is still largely considered as a puzzle (McClennen (1990)). Information decentralization is here offered as a new solution for non-myopic decision making under changing utility of wealth, when change is caused by reference point shifts. Interestingly, this solution goes beyond the two traditional solutions of precommitment and sophisticated choice.

Decentralization resembles precommitment (see, e.g., Schelling (1978) and Elster (1979)), because it extends the scope of the initial viewpoint (and of the resulting preferences) over the future. But it does not imply any kind of regret in the future, nor the canceling of any option. Precommitment alters future alternatives, whereas decentralization stabilizes preferences. In a way, we follow Stigler and Becker (1977), who argue that changing preferences always correspond to a stationary pattern, when the determinants of change are understood. Our work identifies one such determinant (reference point shifts caused by outcome perception) and makes it a controllable variable.

We actually use a "sophisticated" approach as a starting point, i.e., we regard future choices as exogenous constraints for current preferences. But we also assume that the decision maker has initially access to an information system that can determine his future risk attitude. Thus, we end up with a decision procedure where initial choices constrain future preferences. This contrasts with the classical sophisticated solution, which recommends resignation. In that sense, information decentralization in a context of reference point shifts provides an operational device for implementing McClennen's (1990) recent suggestion of "resolute choice" as a solution for planning under changing tastes ("resolute choice" occurs when the future self decides to implement a plan that was elaborated according to initial preferences).

An important aspect of our theory is the distinction between two effects of feedback information: it provokes reference point shifts, and it dictates the implementation of the optimal contingent plan. Information decentralization in a hierarchical context dissociates the

two effects (which would otherwise be simultaneous, for an individual decision maker). This observation allows us to give an original insight about preferences over the timing of uncertainty resolution (in comparison with, e.g., Kreps and Porteus (1978) or Bell (1985)). Indeed, in a hierarchical context, an individual at the top may prefer to delay personal contact with information, while always ordering its immediate processing by a subordinate.

Closely Related Work

Two papers inspired by behavioral decision theory have some similarities with the present essay. Thaler (1985) examines whether a given sequence (x, y) of outcomes gives greater satisfaction when "integrated" (i.e., coded as a sum $x + y$) than when "segregated" (i.e., coded as a sequence of distinct stimuli). In our terms, "integration" corresponds to information decentralization, while "segregation" results from permanently tracking the outcomes. Our approach is different because we analyze the ex ante decision to decentralize information, in a context of uncertainty and learning. Also, we study how a decision to decentralize impacts the chosen policies, an aspect that Thaler ignored.

Kahneman and Lovallo (1993) also discuss loss aversion in an organizational context. But they do not view organizations as systems where preferences are engineered (through controlled perception) so as to achieve some desirable decision behavior. They rather conclude that "the forces that produce risk aversion in experimental studies of individual choice may even be stronger in the managerial context." However, after one of their numerical examples, they mention that "reducing the frequency of evaluations can mitigate the inhibiting effects of loss aversion on risk taking" (p.23), which gives a hint for the present paper.

After having set up the model in the next section, we give a characterization of dynamic choice situations where decentralization of information is optimal (assuming it is freely available). Then, in section III, we perform the reverse analysis to find situations where centralization (i.e., the decision maker permanently keeps track of the outcomes) should be preferred. In section IV, we show that information decentralization constitutes a device to

overcome conservatism, in two important classes of problems. Section V comments on a side effect of information decentralization: it creates a risk of escalation. Section VI concludes.

I. MODEL

A. Dynamic Choice Situation

We consider the problem of a decision maker facing a two-period choice situation.

There is a finite set X of possible monetary outcomes. A representative outcome $x \in X$ is denoted by x_1 when it results from a first period action and by x_2 when it results from a second period action. The set of initially available actions is denoted by A_1 , with representative element a_1 . Each action a_1 is a probability measure defined on a Boolean algebra $\mathcal{B}$ of subsets of X . Both initial action and the obtained first period outcome may determine the performance of period 2 actions (e.g., in case of learning). The choice set in the second period is therefore denoted by $A_2(a_1, x_1)$, with representative element $a_2(\cdot | a_1, x_1)$, a probability measure defined on $\mathcal{B}$.

B. Utility Function

Outcomes are evaluated through an increasing utility function that has the basic features of Kahneman and Tversky's (1979) value function. We assume that accumulated wealth serves as a reference point (as initially suggested by Markowitz (1952)), and the received outcomes are direct stimuli for (dis)satisfaction. We moreover assume loss aversion and a weak form of diminishing sensitivity. These assumptions correspond to realistic features of human "psychophysics of hedonic experience". Empirical support under the von Neumann-Morgenstern framework can be found in Fishburn and Kochenberger (1979), for instance.

Formally, we consider a utility function $v: \bigcup (X, -X, X \oplus X, \{0\}) \rightarrow \mathbb{R}^3$ that is characterized as follows:

$$v(0) = 0, \tag{1}$$

$$v(x) < v(y) \quad \text{for } x < y, \tag{2}$$

³ $-X = \{-x: x \in X\}$ and $X \oplus X = \{x + y: x, y \in X\}$.

$$v(x+y) - v(x) \leq v(y) \quad \text{for } x, y \geq 0 \quad (\text{subadditivity for gains}), \quad (3)$$

$$v(x+y) - v(x) \geq v(y) \quad \text{for } x, y \leq 0 \quad (\text{superadditivity for losses}), \quad (4)$$

$$v(x) < -v(-x) \quad (\text{loss aversion}). \quad (5)$$

In order to focus on the role of loss aversion, we will sometimes work with a special form of v that we call the *constant loss aversion* case:

$$v(x) = \alpha x \text{ if } x \leq 0, v(x) = \beta x \text{ if } x \geq 0, \text{ with } \beta \geq 0 \text{ and } \alpha > \beta.$$

Under constant loss aversion as we define it, (3) and (4) apply with equality signs. We call *degree of loss aversion* the ratio β/α .

Note that v does not change over time (although the implied utility of wealth does).

C. Choice

We assume that the decision maker has free access to a loyal subordinate. Loyalty means that the subordinate is willing to do what he is asked to do. This a cooperative team situation (in the sense of Marschak and Radner (1972)), with no agency problem. *Centralization* occurs when the decision maker does not make use of the subordinate. *Decentralization* occurs when the subordinate is hired initially and the decision maker does not hear about the choice situation until the end of the second period.

Policies are chosen after a dynamic programming procedure that maximizes the sum of discounted expected utilities.

It may seem surprising that we use expected utility as a decision criterion rather than the full prospect theory approach (Tversky and Kahneman (1992)), that is usually associated with reference-dependent utility functions. Our primary reason for doing so is that we focus on those aspects of prospect theory which are not fundamentally incompatible with the von Neumann Morgenstern framework (see, e.g., Markowitz (1952), and Fishburn and Kochenberger (1979)). Our purpose is to bring a dose of descriptive relevance in the classical normative model, and to discuss some implied prescriptions for dynamic choice.⁴

⁴ Kahneman and Lovallo (1993) also use expected utility with reference-dependent utility functions, as an approximation of prospect theory, for the ease of exposition.

We first examine choice *under centralization*. The outcomes are perceived instantaneously when they are generated, and utility for x_1 (x_2) is experienced at the end of the first (second) period. We assume throughout the paper that perception immediately induces the corresponding utility.⁵

At the beginning of second period, expected utility under action $a_2(\cdot|a_1, x_1)$ is thus

$$V_2(a_2(\cdot|a_1, x_1)) = \sum_{x_2 \in X} v(x_2) a_2(x_2|a_1, x_1), \quad (6)$$

and the preferred action is

$$a_2^*(\cdot|a_1, x_1) = \arg \max_{a_2 \in A_2(a_1, x_1)} V_2(a_2(\cdot|a_1, x_1)). \quad (7)$$

Initially, anticipated satisfaction under action a_1 is therefore

$$V_1(a_1) = \sum_{x_1 \in X} [v(x_1) + \delta V_2(a_2^*(\cdot|a_1, x_1))] a_1(x_1). \quad (8)$$

where $\delta \in [0, 1]$ is a discount factor whose interpretation is discussed in the next subsection.

The preferred first period action is

$$a_1^* = \arg \max_{a_1 \in A_1} V_1(a_1). \quad (9)$$

It is apparent that we assume time separability, i.e. a change in x_1 does not impact utility for any x_2 occurring in second period. This is a classical property in dynamic preferences modeling (see, e.g., Meyer (1976)), and a tenet of dynamic programming.⁶

Under decentralization, the generated outcomes are disclosed together at the end of the second period (i.e. exposure to x_1 is delayed, in comparison with the centralization case). Therefore, expected utility associated with action $a_2(\cdot|a_1, x_1)$ takes into account that the

⁵ One might argue that a decision maker should be able to exert willpower so as to postpone experience until the end of the second period, even when earlier perception has occurred. But our work precisely focuses on decision makers who don't have such willpower. Alternatively, our theory can be seen as an analysis of the conditions under which such willpower should prove valuable a priori.

⁶ The descriptive relevance of time separability is discussed in Machina (1989), for instance.

perceived outcome (at the end of period 2) will be the sum of realized x_1 and x_2 .⁷

We have

$$V_2^d(a_2(\cdot|a_1, x_1)) = \sum_{x_2 \in X} v(x_1 + x_2) a_2(x_2|a_1, x_1), \quad (10)$$

and the preferred action is

$$a_2^{*d}(\cdot|a_1, x_1) = \arg \max_{a_2 \in A_2(a_1, x_1)} V_2^d(a_2(\cdot|a_1, x_1)). \quad (11)$$

Initially, the subordinate should pick

$$a_1^{*d} = \arg \max_{a_1 \in A_1} V_1^d(a_1), \quad (12)$$

given that the initially anticipated satisfaction under action a_1 is

$$V_1^d(a_1) = \sum_{x_1 \in X} \delta V_2^d(a_2^{*d}(\cdot|a_1, x_1)) a_1(x_1). \quad (13)$$

D. A Note on the Interpretation of δ ⁸

In our framework, the initial value of a given sequence (x_1, x_2) would be $v(x_1) + \delta v(x_2)$ under centralization, and $\delta v(x_1 + x_2)$ under decentralization. Now, a common rationale for discounting the future (in a context of wealth accumulation) is that an outcome received later has less time to fructify and should therefore be discounted. Why, then, do we apply a different discounting on x_1 under centralization than under decentralization? Indeed, x_1 is received anyway by the agent at the end of the first period (although it may not be perceived by the decision maker), and it can fructify identically under both centralization and decentralization. So, it may appear that our way of applying the discount factor is inconsistent.

⁷ Concerning joint receipt of two outcomes, Luce and Fishburn (1991) have proposed that $v(x_1 \circ x_2) = av(x_1) + bv(x_2) + cv(x_1)v(x_2)$. Such form is incompatible with our assumption that $v(x_1 \circ x_2) = v(x_1 + x_2)$, unless we impose some strong restrictions on v . Tversky and Kahneman (1992) convincingly argue against Luce and Fishburn (and thus in favor of our assumption), on both normative and descriptive grounds.

⁸ I am indebted to Bob Winkler who raised this issue.

However, suppose that x_1 (assuming it is positive) is indeed invested. Such investment produces some additional payoff rx_1 at the end of the second period. Note that *this payoff should be seen as part of x_2* if we conceive that $x_2 = rx_1 + x_2'$, where x_2' is a second period outcome unrelated to the investment of previous gains. This is very much consistent with our notation, where second period lotteries are denoted by $a_2(\cdot|a_1, x_1)$. Now, recall that, under centralization, "the outcomes are perceived instantaneously when they are generated, and utility for x_1 (x_2) is experienced at the end of the first (second) period". Thus, x_1 has a direct impact through $v(x_1)$, plus an indirect impact at the end of the second period, through $v(rx_1 + x_2')$. On the other hand, under decentralization, (x_1, x_2') will generate utility $v(x_1 + rx_1 + x_2')$ at the end of the second period. Thus, in our model, δ is *not* meant to capture the aspect of time preferences that finds its rationale in the investment of early outcome.

But, then, what is the correct interpretation of δ ? It is a subjective discount factor, a *psychological parameter* that reflects impatience towards the experience of positive utilities, and a preference for the postponement of negative utilities.⁹ Since the decision maker's utility at each point in time is generated by exposure to the realized change in wealth, our evaluation functions are fully consistent with this interpretation.

When time preferences are motivated by investment activities only, then δ equals 1.

II. WHEN TO DECENTRALIZE

A. Sufficient Condition for Decentralization

The choice between centralization and decentralization is determined by the sign of $V_1^d(a_1^{*d}) - V_1(a_1^*)$. It is impossible to determine this sign without further specifying the dynamic choice problem. In the present section, we identify several classes of problems that

⁹ A psychological interpretation of discounting is not at all specific to the present work. See, e.g., the textbook by Malinvaud (1972, p.234-35). A particular instance occurs in the theory of repeated games, where the discount factor is often interpreted as a probability of surviving the next period, i.e. the probability that the game will stop between any two consumption dates is $1-\delta$. Although values greater than 1 can be conceived, the traditional notion that decision makers are impatient is kept for the purpose of this paper.

guarantee a preference for decentralization.

Our reasoning is as follows: if the best centralized policy brings more utility under decentralization, then decentralization dominates (i.e. the reference point should be stabilized).

To first illustrate the approach, let us examine the case of risk neutrality (i.e. $v(x) = kx$ for some $k > 0$). Note that there is no issue of stabilizing a reference point here, and decentralization only impacts utility through the discounting factor. Observe that, $\forall a_1 \in A_1$ and $x_1: a_1(x_1) > 0$, we have $a_2^{*d}(\cdot|a_1, x_1) = a_2^*(\cdot|a_1, x_1)$.

Therefore, we can write

$$V_1(a_1^*) = \sum_{x_1 \in X} \left(kx_1 + \delta \sum_{x_2 \in X} kx_2 a_2^*(x_2|a_1^*, x_1) \right) a_1^*(x_1), \quad (14)$$

and

$$\begin{aligned} V_1^d(a_1^*) &= \sum_{x_1 \in X} \left(\delta \sum_{x_2 \in X} k(x_1 + x_2) a_2^*(x_2|a_1^*, x_1) \right) a_1^*(x_1) \\ &= \sum_{x_1 \in X} \left(\delta kx_1 + \delta \sum_{x_2 \in X} kx_2 a_2^*(x_2|a_1^*, x_1) \right) a_1^*(x_1). \end{aligned} \quad (15)$$

Thus, $V_1^d(a_1^{*d}) - V_1(a_1^*) \geq V_1^d(a_1^*) - V_1(a_1^*) = (\delta - 1) \sum_{x_1 \in X} kx_1 a_1^*(x_1)$. This justifies the following

remark:

REMARK. *Under risk neutrality, if the chosen first period action under centralization has negative expected outcome, then decentralization is preferred to centralization.*

Again, note that the advantage of decentralization here merely comes from postponing (expected) bad news.

Now, let us apply the same reasoning to the more realistic non-linear functions introduced in the previous section. Observe that

$$V_1^d(a_1^{*d}) - V_1(a_1^*) \geq$$

$$V_I^d(a_I^*) - V_I(a_I^*) \geq \sum_{x_I \in X} \sum_{x_2 \in X} (\delta v(x_I + x_2) - v(x_I) - \delta v(x_2)) a_2^*(x_2 | a_I^*, x_I) a_I^*(x_I). \quad (16)$$

Now, define the supports

$$X_1 = \{x_1 \in X : a_1^*(x_1) > 0\} \quad (17)$$

$$\text{and } X_2 = \{x_2 \in X : a_2^*(x_2 | a_1^*, x_1) > 0, \text{ for some } x_1 \in X_1\}. \quad (18)$$

We can state:

LEMMA 1. *If, $\forall x_1 \in X_1$ and $\forall x_2 \in X_2$, $c(x_1, x_2) = \delta(v(x_1 + x_2) - v(x_2)) - v(x_1) \geq 0$, then decentralization is preferred to centralization.*

Consider the chosen plan of a centralized decision maker. Because of the shape of his utility function, the decision maker might prefer to perceive any possible first period outcome coupled with the second period outcome (rather than separately), knowing that the associated (dis)satisfaction will be postponed and discounted. When such is the case, decentralization of the centralized plan is beneficial, which is a sufficient (but not necessary) condition to favor decentralization.

B. Under Constant Loss Aversion

For an individual with constant loss aversion, a loss is always more painful than the reduction of a gain by the same amount. Also, because there is no diminishing sensitivity, the utility associated with a loss remains unchanged when that loss is perceived on top of another loss. Moreover, postponement of losses is always welcome.

PROPOSITION 1. *Under constant loss aversion, if the supremum of X_1 is non-positive, then decentralization is preferred to centralization.*

Proof. Consider any $x_1 \in X_1$ and $x_2 \in X_2$. Since $\sup X_1 \leq 0$, we have $x_1 \leq 0$. If $x_2 \leq 0$, then $c(x_1, x_2) = \delta \alpha(x_1 + x_2) - \delta \alpha x_2 - \alpha x_1 = (\delta - 1) \alpha x_1 \geq 0$. If $0 \leq x_2 \leq -x_1$, then $c(x_1, x_2) = \delta \alpha(x_1 + x_2) - \delta \beta x_2 - \alpha x_1 = (\delta - 1) \alpha x_1 + \delta(\alpha - \beta)x_2 \geq 0$. Else, $c(x_1, x_2) = \delta \beta(x_1 + x_2) - \delta \beta x_2 - \alpha x_1 = (\delta \beta - \alpha)x_1 \geq 0$. Thus, the condition expressed in lemma 1 is always satisfied. $\square$

Under constant loss aversion, a plain gain always provides less utility than the reduction of a loss by the same amount. Thus the decision maker can be better off by postponing exposure to a gain in order to reduce a loss in the future. However, postponing exposure to a gain for the sake of transforming it into the reduction of a loss in second period must be traded-off against the negative impact of gain's postponement, due to discounting. Moreover, when the second period loss is small in comparison with the postponed gain, the benefits of postponing exposure are limited. These trade-offs are analyzed by propositions 2 and 3.

PROPOSITION 2. *Under constant loss aversion, if the supremum of X_1 is positive, $\delta \geq \frac{\beta}{\alpha}$, and $\sup X_2 \leq -\sup X_1$, then decentralization is preferred to centralization.*

Proof. Consider any $x_1 \in X_1$ and $x_2 \in X_2$. If $x_1 \leq 0$, the proof of proposition 1 showed that $c(x_1, x_2) \geq 0$. If $x_1 > 0$ and $x_2 \leq -x_1$, then $c(x_1, x_2) = \delta \alpha(x_1 + x_2) - \delta \alpha x_2 - \beta x_1 = (\delta \alpha - \beta)x_1 \geq 0$ if $\delta \geq \frac{\beta}{\alpha}$. Thus the condition expressed in lemma 1 is always satisfied under the above conditions. $\square$

Thus, when the discount factor is higher than the degree of loss aversion, and the second period possible outcomes under the best centralized plan are large losses, then decentralization dominates. When second period possible losses are not so large, we obtain a more restrictive condition on the discount factor.

PROPOSITION 3. *Under constant loss aversion, if the supremum of X_1 is positive, $\sup X_2 \in [-\sup X_1, 0]$, and $\delta \geq b(X_1, X_2) = \frac{\beta \sup X_1}{\beta \sup X_1 + (\beta - \alpha) \sup X_2}$, then decentralization is preferred to centralization.*

Proof. Consider any $x_1 \in X_1$ and $x_2 \in X_2$. If $x_1 \leq 0$, the proof of proposition 1 showed that $c(x_1, x_2) \geq 0$. If $x_1 > 0$ and $-x_1 \leq x_2 < 0$, then $c(x_1, x_2) = \delta \beta(x_1 + x_2) - \delta \alpha x_2 - \beta x_1$, which is non-negative if $\delta \geq b(X_1, X_2)$. If $x_1 > 0$ and $x_2 < -x_1$, then $c(x_1, x_2) = \delta \alpha(x_1 + x_2) - \delta \alpha x_2 - \beta x_1 = (\delta \alpha - \beta)x_1 \geq 0$ if $\delta \geq \frac{\beta}{\alpha}$. Note that $b(X_1, X_2) > \frac{\beta}{\alpha}$ whenever $\sup X_2 \in [-\sup X_1, 0]$. $\square$

C. Under Diminishing Sensitivity

When loss aversion is combined with diminishing sensitivity (i.e. when (3) and (4) apply with inequality signs), our claim that a loss is more painful than the reduction of a gain by the same amount must be qualified. Diminishing sensitivity implies that a large loss generates little dissatisfaction at the margin. Thus, little improvement comes from combining a loss with a small gain. Such little improvement might not compensate for the dropped satisfaction of perceiving the small gain in isolation. Therefore, combining a loss with a small gain is beneficial only when the loss is still sufficiently painful at the margin, i.e. when the loss is not too large. Note that, thanks to diminishing sensitivity, the utility associated with a loss always increases when that loss is cumulated with another loss.

PROPOSITION 4. *Under loss aversion and diminishing sensitivity, if the supremum of X_1 is non-positive and $\inf X_1 \geq -\inf X_2^+$, then decentralization is preferred to centralization.*

Proof. Consider any $x_1 \in X_1$ and $x_2 \in X_2$. Since $\sup X_1 \leq 0$, we have $x_1 \leq 0$. If $x_2 \leq 0$, then, by (4) $c(x_1, x_2) \geq \delta v(x_1) - v(x_1) \geq 0$. If $x_2 \geq -x_1$, replacing $x_1 + x_2$ by $y \geq 0$, and given (3), we have $c(x_1, x_2) = \delta[v(y) - v(y + (-x_1))] - v(x_1) \geq \delta(-v(-x_1)) - v(x_1)$. Therefore, by (5) and given that $x_1 \leq 0$, $c(x_1, x_2) \geq 0$. Thus, when $\inf X_2^+ \geq -\inf X_1$, the condition expressed in lemma 1 is satisfied. $\square$

When discounting is strong enough, the postponement of a first period loss is always attractive, according to proposition 5.

PROPOSITION 5. *Under loss aversion and diminishing sensitivity, if the supremum of X_1 is non-positive and $\delta \leq \min_{x_1 \in X_1} v(x_1)/v(x_1) - v(-x_1)$, then decentralization is preferred to centralization.*

Proof. Consider any $x_1 \in X_1$ and $x_2 \in X_2$. Since $\sup X_1 \leq 0$, we have $x_1 \leq 0$. From the proof of proposition 4, we know that both $x_2 \leq 0$ and $x_2 \geq -x_1$ imply $c(x_1, x_2) \geq 0$. When $0 < x_2 < -x_1$, then $c(x_1, x_2) \geq 0$ if $\delta \leq v(x_1)/v(x_1 + x_2) - v(x_2)$, by definition of $c(x_1, x_2)$.

Note that $v(x_1)/v(x_1 + x_2) - v(x_2) > v(x_1)/v(x_1) - v(-x_1)$, $\forall x_2 \in]0, -x_1[$. Thus, when

$\delta \leq v(x_1)/v(x_1) - v(-x_1) \quad \forall x_1 \in X_1$, the condition expressed in lemma 1 is satisfied. $\square$

Diminishing sensitivity also imposes a qualification of proposition 2. Indeed, the reduction of a loss is not always more enjoyable than a plain gain by the same amount, especially if the loss is large and creates little dissatisfaction at the margin. Also, diminishing sensitivity implies that a gain yields less satisfaction when perceived on top of another gain than when perceived in isolation.

PROPOSITION 6. *Under loss aversion and diminishing sensitivity, if the supremum of X_1 is positive, $\delta = 1$ and $X_2 \subset [-\inf X_1^+, 0]$, then decentralization is preferred to centralization.*

Proof. Consider any $x_1 \in X_1$ and $x_2 \in X_2 \subset [-\inf X_1^+, 0]$. For $x_1 \leq 0$, $c(x_1, x_2) \geq 0$, by proposition 4. (Note that if $x_1 > 0$ and $x_2 > 0$, then $c(x_1, x_2) < 0$, by (3)). If $x_1 > 0$ and $x_2 = 0$, then $c(x_1, x_2) = 0$. If $x_1 > 0$ and $x_2 = -x_1$, then $c(x_1, x_2) = -v(-x_1) - v(x_1) > 0$. When $-x_1 < x_2 < 0$, replacing $x_1 + x_2$ by $y > 0$, and by (3) and (5), we can write $c(x_1, x_2) = v(y) - v(x_2) - v(y + (-x_2)) \geq -v(-x_2) - v(x_2) \geq 0$. Thus, if $x_2 \in [-\inf X_1^+, 0] \quad \forall x_2 \in X_2$, then the condition expressed in the lemma is satisfied. $\square$

III. WHEN TO CENTRALIZE

If the best decentralized policy brings more expected utility when centralized, then centralization dominates, i.e. shifts of the reference point should be allowed.

Observe that

$$V_1^d(a_1^{*d}) - V_1(a_1^*) \leq \sum_{x_1 \in X} \sum_{x_2 \in X} (\delta v(x_1 + x_2) - v(x_1) - \delta v(x_2)) a_2^{*d}(x_2 | a_1^{*d}, x_1) a_1^{*d}(x_1). \quad (19)$$

Now, define the supports

$$X_1^d = \{x_1 \in X : a_1^{*d}(x_1) > 0\} \quad (20)$$

$$\text{and } X_2^d = \{x_2 \in X : a_2^{*d}(x_2 | a_1^{*d}, x_1) > 0, \text{ for some } x_1 \in X_1^d\}. \quad (21)$$

We can state:

LEMMA 2. If, $\forall x_1 \in X_1^d$ and $\forall x_2 \in X_2^d$, $c(x_1, x_2) = \delta(v(x_1 + x_2) - v(x_2)) - v(x_1) \leq 0$, then centralization is preferred to decentralization.

Decentralization postpones exposure to first period gains. Under discounting, this is considered a drawback.

PROPOSITION 7. Under constant loss aversion, if the infimum of X_1^d is non-negative and $\delta \leq \frac{\beta}{\alpha}$, then centralization is preferred to decentralization.

Proof. Consider any $x_1 \in X_1^d$ and $x_2 \in X_2^d$. If $x_2 \geq 0$, then $c(x_1, x_2) = \delta \beta(x_1 + x_2) - \delta \beta x_2 - \beta x_1 = (\delta - 1) \beta x_1 \leq 0$. If $x_2 < -x_1$, then $c(x_1, x_2) = \delta \alpha(x_1 + x_2) - \delta \alpha x_2 - \beta x_1 = (\delta \alpha - \beta) x_1$, which is negative if $\delta \leq \frac{\beta}{\alpha}$. If $-x_1 \leq x_2 < 0$, then $c(x_1, x_2) = \delta \beta(x_1 + x_2) - \delta \alpha x_2 - \beta x_1$ is negative if and only if $\delta \leq \frac{\beta x_1}{\beta x_1 + (\beta - \alpha) x_2}$. Note that $-x_1 \leq x_2 < 0$ implies $\frac{\beta}{\alpha} \leq \frac{\beta x_1}{\beta x_1 + (\beta - \alpha) x_2}$, and therefore $\delta \leq \frac{\beta}{\alpha}$ is sufficient in this case too. $\square$

Centralization also dominates under diminishing sensitivity, when the best decentralized policy leads to a sequence of gains.

PROPOSITION 8. Under loss aversion and diminishing sensitivity, if the infimum of both X_1^d and X_2^d is positive, then centralization is preferred to decentralization.

Proof. By (3). $\square$

Propositions 7 and 8 can trivially be used to generate necessary conditions for information decentralization.

IV. DECENTRALIZATION AS AN ANTI-CONSERVATISM DEVICE

In section 2, we have characterized dynamic choice situations where a decision maker would benefit from information decentralization. Now, what can we say about the resulting (controlled) risk attitude? Is decentralized decision making more risk-taking when it arises? In this section, we introduce two classes of dynamic situations where the opportunity to decentralize information enhances risk taking.

A. Entry Cost

Suppose that the decision maker faces a choice between the status quo (no loss, no gain over the two periods) and a risky action $a(\cdot)$ that has positive expected payoff, but is only accessible after a first period sure loss ℓ . Such risky plan with entry cost could be a project in research and development, for instance.

Note that if (i) the decision maker has constant loss aversion or (ii) the decision maker is characterized by both loss aversion and diminishing sensitivity with

$$\text{either } \ell \geq -\inf \{x \in X^+ : a(x) > 0\} \text{ or } \delta \leq v(\ell)/v(\ell) - v(-\ell),^{10} \quad (22)$$

then

$$\delta(v(\ell + x_2) - v(x_2)) - v(\ell) \geq 0 \quad \forall x_2 : a(x_2) > 0, \quad (23)$$

as shown in the proofs of propositions 1, 4 and 5. Thus, the risky plan yields higher anticipated expected utility under information decentralization than under centralization. Access to a hierarchical system renders choice of the risky plan more likely, because it softens the impact of the entry cost for a large class of decision makers.

B. Repeated Situations

Now, assume that the decision maker must choose, at the beginning of each period, between the status quo and a given alternative $a(\cdot)$ that has a negative expected utility but a positive expected value. One can think of a decision whether to open a shop everyday in a stationary environment, for instance.

Solving the dynamic problem backward, we first observe that a centralized decision maker, at the beginning of the second period, would prefer the status quo over $a(\cdot)$. Initially, his choice thus reduces to $a(\cdot)$ played once versus the status quo, and he would stick with the status quo.

Under decentralization, the second period decision depends on the first period outcome. If $a(\cdot)$ has been picked initially, then the status quo is selected in second period if and only if the realized outcome x_1 was such that it is *a priori* beneficial to combine x_1 with $a(\cdot)$. In other

¹⁰ i.e., the initial loss is not too large.

words, status quo is selected if and only if

$$v(x_1) > \sum_{x_2 \in X} v(x_1 + x_2) a(x_2). \quad (24)$$

Thus, from the initial standpoint, information decentralization implies that any policy starting with $a(\cdot)$ has an anticipated value that is at least as great as the one associated with $a(\cdot)$ played once in second period. Indeed, by departing from the status quo initially, a decentralized decision maker anticipates

$$\sum_{x_1 \in X} \delta \max \left[v(x_1), \sum_{x_2 \in X} v(x_1 + x_2) a(x_2) \right] a(x_1) \quad (25)$$

If the above value is positive, then information decentralization leads the decision maker to take the same initial risk as a risk-neutral individual would take.

V. ESCALATION AS A SIDE EFFECT

A possible side effect of information decentralization is escalation. Escalation is frequently observed in organizational behavior (see, e.g., Staw and Ross (1989)). It occurs when a decision maker fails to terminate a project that proves unsuccessful. It generally results from a failure to consider sunk costs as really sunk. Because of diminishing sensitivity (which implies risk-seeking in the zone of losses) a decision maker "who has not made peace with his losses is likely to accept gambles that would be unacceptable for him otherwise" (Kahneman and Tversky (1979), p. 287).

Under information decentralization, the decision maker commits not to get exposure to intermediary payoffs, and costs (losses) that are really sunk after the first period are not psychologically so. By stabilizing his reference point, the decision maker prevents himself to "make peace" with intermediary payoffs, and this may naturally lead to escalating policies.

Escalation and the underlying sunk cost effect are generally viewed as irrational. On the other hand, we have shown that information decentralization is often desirable and provides a mechanism to overcome loss aversion and resulting conservatism. Thus, we argue that escalation can often be seen as a side effect of an otherwise desirable and rational decision making process. The meta-decision about whether to postpone reports on intermediary

outcomes often deals with a trade-off between a chance to avoid the status quo bias on the one hand, and a risk of escalation on the other hand.

VI. CONCLUSION

This paper has shown how, by allowing contingent policy implementation while maintaining the decision maker ignorant of the current decision being taken, a hierarchical device results in a transformation of individual risk preferences. This transformation can be anticipatively attractive for a decision maker.

We have not taken any normative position about the rationality of the features that we examined (loss aversion, reference point shifts, and the associated conservatism). We tend to respect them as real determinants of a decision maker's degree of satisfaction. However, as we have seen, the very target of improving utility may require a transformation of such features. In other words, the examined biases are sometimes self-defeating: a decision maker with reference point shifts may sometimes find it valuable to fight against such shifts. This aspect of our work is probably of particular interest to those economists who are concerned about whether individual biases survive on the market place (see, e.g., the interesting discussion by Frey and Eichenberger (1994)). Overall, (and even though we only focused on a few of the most descriptively relevant features of human decision behavior) our work qualifies as prescriptive.

Our results also have an explanatory potential. We offered insights about why hierarchies exist and about how individual tendencies get transformed in the social context.

We found, for instance, that the possibility of initial losses (provided they cannot be too large) constitutes an incentive to switch from information centralization to information decentralization (proposition 4). Under constant loss aversion, information decentralization is always preferred, provided impatience is not too strong (propositions 1, 2 and 3). Moreover, access to a subordinate can help overcome conservatism in the presence of entry costs, as well as in repeated choice contexts. Escalation has been re-interpreted as a side effect of an otherwise desirable feedback system.

We have made rather specific assumptions regarding the utility of a sequence of outcomes:

we assumed an additive multi-attribute utility function that favors first period utility. Other assumptions could be examined, but the basic contribution of this paper would remain unchanged: when experienced utility depends on the frequency of exposure to the realized outcomes, a decision to decentralize information determines the chosen policies. Whether our set of assumptions is relevant is an empirical issue. Laboratory experiments should be conducted to see whether actual preferences correspond to the predictions made in section II to IV. At least, our general insight that decentralization renders policies with possible intermediary losses more attractive corresponds well to Chandler's (1962) notion that decentralization is especially necessary when the environment is such that growth is not guaranteed.

Other assumptions about the nature and dynamics of reference points might also be examined. We assumed that the reference point is the current known wealth position and that the decision maker gets immediately accustomed to his/her new wealth level, when it is disclosed. Reference points can also be expectation/aspiration levels, and reference point updating could be slower than assumed here (see, e.g., Wathieu (1994)). The impact of information decentralization on preferences under these assumptions remains an open question.

Future research might also relax the assumption of a free and loyal subordinate. Indeed, the decision maker's preferences over the timing of communication in the hierarchy might conflict with the best interest of the subordinate, for some reason.

REFERENCES

- Arkes, Hal R. (1991). 'Costs and benefits of judgment errors: implications for debiasing,' *Psychological Bulletin*, vol. 110, pp. 486-498.
- Barnard, Chester (1968). *The Functions of the Executive*. Cambridge: Harvard University Press.
- Bell, David E. (1985). 'Disappointment in decision making under uncertainty,' *Operations Research*, vol. 30, pp. 961-981.
- Chandler, Alfred D. Jr. (1962). *Strategy and Structure*. Cambridge: MIT Press.
- Elster, Jon (1979). *Ulysses and The Sirens: Studies in Rationality and Irrationality*. Cambridge: Cambridge University Press.
- Fishburn, Peter C. and Kochenberger, Gary A. (1979). 'Two-piece Von Neumann-Morgenstern utility functions,' *Decision Sciences*, vol. 10, pp. 503-518.
- Fishburn, Peter C. (1988). *Nonlinear Preference and Utility Theory*. Baltimore: John Hopkins University Press.
- Frey, Bruno S. and Eichenberger, Reiner (1994). 'Economic incentives transform psychological anomalies,' *Journal of Economic Behavior and Organization*, vol. 23, pp. 215-234.
- Geanakoplos, John and Milgrom, Paul (1991). 'A theory of hierarchies based on limited managerial attention,' *Journal of the Japanese and International Economies*, vol. 5, pp. 205-225.
- Kahneman, Daniel and Tversky, Amos (1979). 'Prospect theory: an analysis of decision under risk,' *Econometrica*, vol. 47, pp. 263-291.
- Kahneman, Daniel and Lovallo, Dan (1993). 'Timid choices and bold forecasts: a cognitive perspective on risk taking,' *Management Science*, vol. 39, pp. 17-31.
- Kreps, David M. and Porteus, Evan L. (1978). 'Temporal resolution of uncertainty and dynamic choice theory,' *Econometrica*, vol. 46, pp. 185-200.
- Loewenstein, George F. (1988). 'Frames of mind in intertemporal choice,' *Management Science*, vol. 34, pp. 200-214.
- Luce, R. Duncan and Fishburn, Peter C. (1991). 'Rank- and sign- dependent linear utility

- models for finite first-order gambles,' *Journal of Risk and Uncertainty*, vol. 4, pp. 29-59.
- Machina, Mark J. (1989). 'Dynamic consistency and non-expected utility models of choice under uncertainty,' *Journal of Economic Literature*, vol. 27, pp. 1622-1668.
- Malinvaud, Edmond (1972). *Lectures on Microeconomic Theory*. Amsterdam: North-Holland.
- Markowitz, Harry (1952). 'The utility of wealth,' *Journal of Political Economy*, vol. 60, pp. 151-158.
- Marschak, Jacob and Radner, Roy (1972). *Economic Theory of Teams*. New Haven: Yale University Press.
- McClennen, Edward F. (1990). *Rationality and Dynamic Choice: Foundational Explorations*. Cambridge: Cambridge University Press.
- McCrimmon, Kenneth R. and Wehrung, Donald A. (1986). *Taking Risks*. New York: The Free Press.
- Meyer, Richard F. (1976). 'Preferences over time.' In Ralph L. Keeney and Howard Raiffa, *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. New York: John Wiley.
- Postrel, Steven and Rumelt, Richard P. (1992). 'Incentives, routines, and self-command,' *Industrial and Corporate Change*, vol. 1, pp. 397-425.
- Radner, Roy (1992). 'Hierarchy: the economics of managing,' *Journal of Economic Literature*, vol. 30, pp. 1382-1415.
- Samuelson, William and Zeckhauser, Richard (1988). 'Status quo bias in decision making,' *Journal of Risk and Uncertainty*, vol. 1, pp. 7-59.
- Schelling, Thomas C. (1978). 'Economics, or the art of self-management,' *American Economic Review*, vol. 68, pp. 290-294.
- Staw, Barry M. and Ross, Jerry (1989). 'Understanding behavior in escalation situations,' *Science*, vol. 246, pp. 216-220.
- Stigler, George and Becker, Gary (1977). 'De gustibus non est disputandum,' *American Economic Review*, vol. 67, pp. 76-90.
- Strotz, R. H. (1956). 'Myopia and inconsistency in dynamic utility maximization,' *Review of*

Economic Studies, vol. 23, pp. 165-180.

Thaler, Richard H. (1985). 'Mental accounting and consumer choice,' *Marketing Science*, vol. 4, pp. 199-214.

Tversky, Amos and Kahneman, Daniel (1991). 'Loss aversion in riskless choice: a reference-dependent model,' *Quarterly Journal of Economics*, vol. 106, pp. 1039-1061.

Tversky, Amos and Kahneman, Daniel (1992). 'Advances in prospect theory: cumulative representation of uncertainty,' *Journal of Risk and Uncertainty*, vol. 5, pp. 297-323.

Wathieu, Luc (1994). *Reference Points and Loss Aversion: Dynamic and Organizational Implications*. Fontainebleau: INSEAD (Unpublished Dissertation).

Please forward your requests for working papers to the following address:

Executive Officer
Department of Marketing
School of Business and Management
Hong Kong University of Science & Technology
Clear Water Bay
Kowloon, Hong Kong

www: <http://www.bm.ust.hk/~mark/>
Enquiries: Tel (852) 2358 7700 Fax (852) 2358 2429