Macroeconomic Dynamics, **6**, 2002, 5–18. Printed in the United States of America.
DOI: 10.1017.S1365100501010185

ARTICLES

RISK SENSITIVITY, A STRANGELY PERVASIVE CONCEPT

PETER WHITTLE
University of Cambridge

The theme of this paper is that the simple concept of risk sensitivity raises ideas, shows associations, and forces clarification of issues that one would have thought quite unrelated.

Keywords: Risk, Robustness, Large Deviation Theory

1. THE PROPERTY OF RISK SENSITIVITY

Risk sensitivity is related to the idea of *utility*. That is, the value to an individual of a given sum of money r may not be proportional to r itself, but may be some more general function $U(r)$, known as the *utility function*. U usually will be an increasing function of r, so that saying that one would like $U(r)$ to be large is the same as saying that one would like r to be large. However, the rescaling implied by a utility function has an effect when one comes to consider the trade-off between competing activities in which one might invest. More than that, the utility function is essentially *defined* by the requirement that the utility of an *uncertain* return r should be just $EU(r)$, where E indicates an expectation over the random variable r.

If the function $U(r)$ is concave (see Figure 1), then $EU(r) < U[E(r)]$. That is, for a given expected return, the individual always prefers a *certain* return. He is then said to be *risk averse*. In the contrary case, when $U(r)$ is convex, the opposite inequality holds, and the individual likes a degree of uncertainty. He is then said to be *risk seeking*. In the transitional case, when $U(r)$ is linear, he is said to be *risk neutral*.

Risk sensitivity has immediate implications. For example, if one must hold an investment portfolio fixed over a period of time, then a risk-seeking individual will concentrate on the single most attractive investment, whereas a risk-averse individual will spread his investment, thus trading peak performance for assured performance.

One could equally well say that a risk-seeking (risk-averse) individual is optimistic (pessimistic), since he implicitly assumes that uncertainties will turn out to his advantage (disadvantage). This view will emerge more clearly in the next section.

Address correspondence to: Professor Peter Whittle, Statistical Laboratory, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WB, UK; e-mail: secretary@statslab.cam.ac.uk.

© 2002 Cambridge University Press 1365-1005/02 $9.50

https://doi.org/10.1017/S1365100502027025 Published online by Cambridge University Press

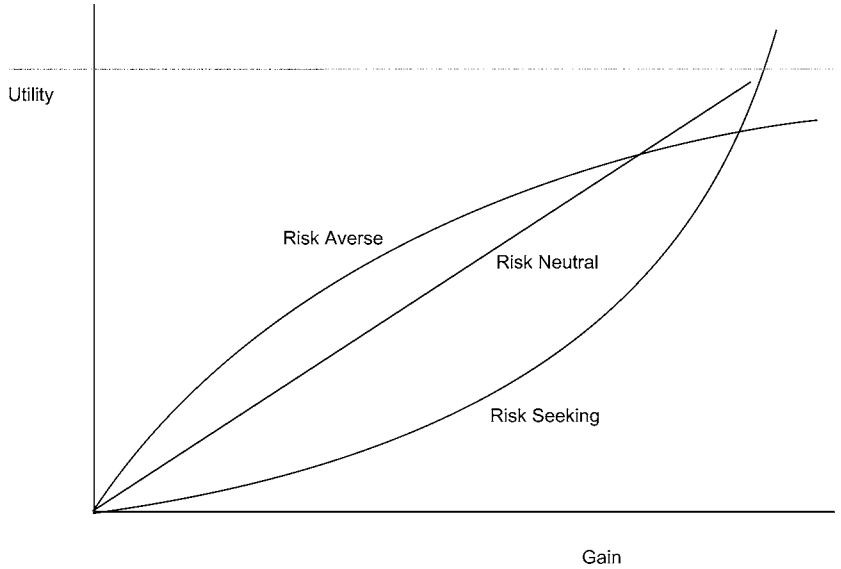


FIGURE 1. Convex (RS), concave (RA), and linear (RN) utility functions.

One can phrase a decision problem as a question of either maximizing returns, in some sense, or of minimizing costs. The conventional view in control theory is that of minimizing costs, and it is then easier for some purposes to take that convention. So, instead of speaking of the utility $U(r)$ of a return r, we shall speak of the *disutility* $L(c)$ of a cost c. The rough connection is that $L(c) = -U(-c)$, and so it is now concave and convex L that correspond to risk-seeking and risk-averse behavior, respectively.

It is also helpful sometimes to invert the L transformation after having taken expectations and so return to a cost scale. Thus,

$$\gamma = L^{-1}[EL(c)] \tag{1}$$

is the fixed cost which is equivalent to an uncertain cost c. Here, L^{-1} is the function inverse to L, and definition (1) works whether L is increasing or decreasing. That is, if one wants costs to be small, then one wants γ to be small, whether L is increasing or decreasing.

One disutility function that is of special interest is the exponential; $L(c) = e^{-\theta c}$, where the parameter θ measures the degree and nature of risk sensitivity. The exponential is always convex, but one wishes to maximize or minimize $e^{-\theta c}$ according to whether θ is positive or negative. This implies, if one thinks about it, that the individual is risk seeking or risk averse according to whether θ is positive or negative. In either case, one will wish to minimize the effective cost

$$\gamma = -\theta^{-1}\log[Ee^{-\theta c}].$$

https://doi.org/10.1017/S1365100502027025 Published online by Cambridge University Press

The exponential disutility is special in that it shows a constant cost elasticity; the proportional change in disutility $\delta L/L$ for a small change δc in c is $-\theta\delta c$, independent of c. However, it is special in other ways, which we list below, with explanations of some of the terms following immediately thereafter:

(i) The transformation approximately decouples expectation and variability, in that

$$\gamma \approx E(c) - \frac{1}{2}\theta \mathrm{var}(c)$$

for small $\theta \mathrm{var}(c)$.

(ii) The parameter θ is a natural risk sensitivity parameter, in that the optimizer's attitude ranges from extreme optimism to extreme pessimism as θ decreases from large positive values to large negative ones.

(iii) It leads to considerations of robustness, and in particular links up with H_∞ theory.

(iv) It shows a mathematical inevitability, in that it fits in naturally with LQG theory.

(v) It also fits in naturally with large-deviation theory in the non-LQG case.

Here LQG theory is the decision theory of models with linear relations, quadratic costs, and Gaussian (normal) random disturbances. These cases permit a very complete and simple treatment; see Sections 2, 3, and 6. Large-deviation theory is a powerful formalism for models that are close to deterministic, in a particular sense. It supplies a kind of equivalent of LQG theory for much more general cases; see Section 8. We touch on H_∞ theory in Section 4.

The risk-sensitive approach is worth following up for its own rationale: that in characterizing performance of a policy it gives weight to variability as well as to expectation of reward. It does show surprising connections, however, and links in particular with concepts of robustness (i.e., of insensitivity to errors in model specification).

2. A SIMPLE LQG EXAMPLE; THE PHANTOM OTHER; OPTIMISM, PESSIMISM, AND NEUROSIS

Let us take the simplest one-stage example. Suppose that x is the distance of an object from its desired position, and that the application of a correction u will bring it to $x - u$. Suppose that the cost of this maneuver, if one leaves matters there, is

$$\mathbf{C} = \frac{1}{2}[Qu^2 + R(x-u)^2].$$

Here, the two terms represent, respectively, the cost of correction (i.e., of control) and of final displacement from the desired position. One finds easily that the optimal value of u and the minimal cost are

$$u = \frac{Rx}{Q+R}, \qquad F(x) = \frac{1}{2}\frac{QRx^2}{Q+R}. \tag{2}$$

A realistic modification is to suppose that the shift achieved is somewhat unpredictable, so that $x - u$ is replaced by $x - u + \varepsilon$. Here the disturbance ε is supposed

https://doi.org/10.1017/S1365100502027025 Published online by Cambridge University Press

normally distributed with zero mean and variance N, and with value unknown at the time u is to be determined. The cost function then becomes

$$\mathbf{C} = \frac{1}{2}[Qu^2 + R(x - u + \varepsilon)^2] \tag{3}$$

One finds that the optimal control still has the previously determined value (2) (a very special case of the property of *certainty equivalence*) and that the effect on the minimal cost $F(x)$ is simply to add a term $\frac{1}{2}RN$.

Now consider a risk-sensitive version of the problem, in which u is chosen to extremize

$$E(e^{-\theta \mathbf{C}}) \propto \int e^{-\theta \mathbf{C} - \mathbf{D}}\, d\varepsilon,$$

where $\mathbf{D} = \frac{1}{2}\varepsilon^2/N$.

Specifically, we should choose u to maximize or minimize this form, according to whether θ is positive or negative. The exponent is a quadratic form in the variables x, u, and ε, and so, one can expect evaluations to follow by standard methods. In fact, the natural course is to write the exponent as $-\theta \mathbf{S}$, where

$$\mathbf{S} = \mathbf{C} + \theta^{-1}\mathbf{D} \tag{4}$$

is the *stress*. Then we find that the rule for determining the optimal control u is to choose u to minimize $\mathbf{S}$, with ε being chosen to minimize or maximize $\mathbf{S}$ according as θ is positive or negative. (In the sequel we use the term *extremize* to mean exactly this.)

This implies a very interesting view of the situation. For our simple example, the stress is

$$\mathbf{S} = \frac{1}{2}[Qu^2 + R(x - u + \varepsilon)^2 + \varepsilon^2/\theta N]. \tag{5}$$

The random disturbance thus appears as just another control term like u, carrying a cost $\frac{1}{2}\varepsilon^2/\theta N$ just as the control u carried a cost $\frac{1}{2}Qu^2$. It is then as though there were another agent, the "phantom other" who exerts the control ε at the same time as the optimizer exerts the control u. If θ is positive (the risk-seeking case), then the phantom other is helping the optimizer, in that they both agree that the stress is to be minimized. However, if θ is negative (the risk-averse case), then the phantom other is opposing the optimizer, in that he is trying to *maximize* stress.

So, what started out as a one-person game has turned into a two-person game. We see now the appropriateness of characterizing risk-seeking and risk-averse attitudes as optimistic and pessimistic, respectively. The two-person view implies that the uncertain element ε takes values that are to the optimizer's advantage or disadvantage, according to whether θ is positive or negative—just what one means by optimism or pessimism.

These ideas were developed first by Jacobson (1973, 1977) in one of the first studies of optimization under a risk-sensitive criterion.

https://doi.org/10.1017/S1365100502027025 Published online by Cambridge University Press

One finds that the optimal value of control and the extremal value of stress are

$$u = \frac{Rx}{Q + R + \theta QRN}, \qquad F(x) = \frac{1}{2}\frac{QRx^2}{Q + R + \theta QRN} \tag{6}$$

In fact, $F(x)$ differs from the equivalent fixed cost γ by a constant term. The value determined for ε is just $-\theta QNu$.

Note that expressions (6) both become infinite as θ decreases through the critical value:

$$\theta_{\text{crit}} = -\frac{1}{N}\left(\frac{1}{Q} + \frac{1}{R}\right).$$

This value is negative and marks a point at which the optimizer is so pessimistic that his apprehension of uncertainties completely overrides the assurance given by known statistical behavior. It is not stretching matters to term this the point of *neurotic breakdown*. There can be a corresponding optimistic extreme, *euphoria*, if the cost function contains quadratic *reward* terms.

3. A MORE GENERAL LQG EXAMPLE

The calculations of the last section can all be carried through for the general LQG model, which can be characterized simply as a model in which (i) the process variable x, the control u, and the observation y are all vector-valued, permitted to take any value in their respective spaces; (ii) the components **C** and **D** of stress are both quadratic in the variables; and (iii) the stretch of time over which operation is to be optimized is fixed.

The one-stage example of the last section did not bring out the role of time. The simplest multistage model of some generality has a cost function of the form

$$\mathbf{C} = \frac{1}{2}\sum_{t=0}^{T-1}\left[x_t^T R x_t + u_t^T Q u_t\right] + \frac{1}{2}(x^T \Pi x)_T \tag{7}$$

and a *plant equation*

$$x_t = Ax_{t-1} + Bu_{t-1} + \varepsilon_t, \tag{8}$$

where x_t is the value of the process variable (in fact, a *state variable* in this case) at time t, etc. The quadratic forms in (7) penalize deviations of x and u from zero at all times, and so, this is a problem of *regulation* to $(0, 0)$. T is the *horizon*, the point at which optimization ceases. The cost function (7) is a sum of *instantaneous costs*, incurred at the times $t < T$, and a *terminal cost*, incurred at the horizon point T.

The only stochastic element in the model as it stands is the *plant noise* $\{\varepsilon_t\}$, whose elements we suppose to be normally and independently distributed with zero mean and covariance matrix N. The term **D** then has the form

https://doi.org/10.1017/S1365100502027025 Published online by Cambridge University Press

$$\mathbf{D} = \frac{1}{2}\sum_{t=1}^{T} \varepsilon_t^T N^{-1} \varepsilon_t, \tag{9}$$

where ε_t is to be identified with $x_t - Ax_{t-1} - Bu_{t-1}$. We know that $\mathbf{C}$ is the cost, and can well term $\mathbf{D}$ the *discrepancy*, since it is indeed a measure of the inexactness of the relationship between x_t and past x and u.

The derivation of the optimal control $\{u_t\}$ in the risk-neutral case is well known, and calculations can be given a nice recursive form. [See, e.g., Bertsekas (1987) or Whittle (1996b).] This all goes through for the risk-sensitive case, following the pattern of the last section. Stress is defined by (4) as before, and the optimizing principle again turns out to be: At a given time t, *minimize* with respect to all currently undetermined decisions and *extremize* with respect to all current unobservables. The recursive form of the risk-neutral calculations carries over.

This was very much the point that Jacobson reached in his 1977 paper. What defeated him and others for some time was the treatment of the case with incomplete observation of the state variable x. The treatment in the risk-neutral case is well established: One derives an estimate $\hat{x}_t$ of x_t based on information at time t from the Kalman filter and simply substitutes this for x_t in the optimal deterministic control rule. By the certainty equivalence rule valid for LQG models, this yields the optimal control. However, the simple recursivity of the calculations seemed to fail in the risk-sensitive case. Essential changes in understanding are required, to which we return in Section 6.

Explicit results are simpler for the continuous-time version of the problem, when the plant equation (8) becomes

$$\frac{dx}{dt} = Ax + Bu + \varepsilon \tag{10}$$

and the sum over time in the cost function (7) is replaced by a corresponding integral. For simplicity, let us consider the infinite-horizon version of the scalar case. Then, the optimal control is

$$u = -(B\Pi/Q)x, \tag{11}$$

where Π is a solution of the quadratic $R + 2A\Pi - J\Pi^2 = 0$ and

$$J = (B^2/Q) + \theta N. \tag{12}$$

The term B^2/Q is a cost-effective measure of control power, in that it is the ratio of control effectiveness to control cost. We see that to this is added a scalar multiple θN of *noise power*. This is again an indication that it is as though random variation helps control in the optimistic case, when θ is positive, and hinders it in the pessimistic case, when θ is negative. One finds that the critical value of θ is

https://doi.org/10.1017/S1365100502027025 Published online by Cambridge University Press

$$\theta_{\text{crit}} = \begin{cases} -\dfrac{B^2}{NQ} & \text{if} \quad A \geq 0 \\ -\dfrac{B^2}{NQ} - \dfrac{A^2}{NR} & \text{if} \quad A \leq 0 \end{cases} . \tag{13}$$

The cases $A > 0$ and $A < 0$ are just those in which the uncontrolled system is unstable or stable, respectively.

In optimizing with respect to decisions yet unmade, one determines the optimal current decision; how is one to regard determinations of future decisions obtained in this way? They are in fact best estimates of what optimal future decisions will be, and so constitute a provisional forward plan. To have this set out explicitly can be a valuable support to intuition, even if this does demand rather more calculation than is required by the recursive reduction of the problem.

The full effects of the shift in view from a risk-neutral criterion to a risk-sensitive one are best summarized when, in Section 6, we consider the case in which the state variable is imperfectly observed.

4. CONNECTION WITH THE H_∞ CRITERION

The risk-sensitive criterion is intrinsically a statistical one; if there is no uncertainty, there is no risk. The H_∞ criterion is intrinsically a deterministic one; it optimizes the system against the most destabilizing deterministic disturbance that could be envisaged within a certain class. Yet, remarkably, the two meet up, as was demonstrated by Glover and Doyle (1988). They meet up for LQG models in the infinite-horizon situation, in which optimal decision rules are independent of time. It turns out then that the optimal policy on the H_∞ criterion is just the optimal policy on the risk-sensitive criterion when θ takes its critical value, that is, when the optimizer is just on the verge of hopeless pessimism.

Pessimism is, of course, the common feature of the two situations: One looks for a worst-case deterministic input in one characterization, and for the worst that a stochastic input can throw at one in the other. Nevertheless, the bridging was unexpected.

More explicitly, the connection is as follows: One can define the inputs to the model as those disturbances to which it is subject, deterministic or stochastic, respectively. One can define the outputs as those quantities (such as control effort and deviations from target) that one would like to keep small, and which are penalized in the cost function. If the model and the control rule are linear, then output is related to input by a *transfer function* $G(i\omega)$, a matrix that relates output to input at frequency ω. Define $\sigma^2(G)$ as the maximum over real ω of the maximum eigenvalue of the square matrix $G(i\omega)G^T(-i\omega)$, which we write as $G\bar{G}$ for brevity. Then, $\sigma^2(G)$ is the maximal possible ratio of mean square output to mean square input, and, in this sense, characterizes worst-case performance. The quantity $\sigma(G)$, sometimes written $\|G\|_\infty$, is known as the H_∞ *norm* of G, and on the H_∞ criterion

https://doi.org/10.1017/S1365100502027025 Published online by Cambridge University Press

one seeks the G that minimizes this. That is, if one proposes a stabilizing linear control rule, then this determines a transfer function G, and it is the control rule that one varies to minimize $\sigma(G)$.

Turning now to the risk-sensitive criterion, it was shown by Glover and Doyle that the expected cost per unit time under a policy that yields the normalized transfer function G is

$$\Gamma = \lim_{T\to\infty}(\gamma/T) = \frac{1}{4\pi\theta}\int \log[|I + \theta G\bar{G}|]\,dw. \qquad (\mathbf{14})$$

However, the greatest value of θ that makes the matrix singular for some ω, and so marks neurotic breakdown, is just $\theta = -[\sigma^2(G)]^{-1}$ and, in minimizing $\sigma(G)$ with respect to policy, one also minimizes the breakdown value of θ. The H_∞-optimal policy is hence just the risk-sensitive policy that is optimal at the breakdown point.

It is this bridging that indicates that the risk-sensitive control may also have robustness properties; see Section 7.

5. IMPLICATIONS FOR DISCOUNTING

Let us take the infinite-horizon version of the model of Section 2 for simplicity, and write the components (7) and (9) of stress as $\mathbf{C} = \sum_t c_t$ and $\mathbf{D} = \sum_t v_t$. Now, it is common to consider a *discounted* cost function, for which it is assumed that $\mathbf{C}$ takes the form $\sum_t \beta^t c_t$. Here, β is the *discount factor*, the reciprocal of the factor by which capital would grow over unit time by compound interest. However, when one performs a risk-sensitive optimization on this assumption, one finds that the optimal control seems to be time dependent, in fact converging to the optimal risk-neutral control with increasing time.

This seems quite wrong—one is surely facing the same situation at all stages. The formal reason why it comes about is because the terms in $\mathbf{C}$ are discounted, whereas those in $\mathbf{D}$ are not. One is then impelled to take a second look at the concept of discounting.

One returns to basics by recognizing that, in running an enterprise and using discounted criteria, one effectively has money (or debts) in two pockets. By the side of the enterprise itself, one has a financial float, a bank account in which money can be invested (or overdrawn), earning interest at a rate corresponding to the growth factor, β^{-1}. Cash flows back and forth between enterprise and account. At any time, one has assets (possibly negative) in both, and the distribution of these assets is determined by the policy by which one runs the enterprise. So, there really are two components of state: the state variable x of the enterprise and the bank balance, which we shall denote as w.

It then turns out [see Whittle (1996a) or (1996b)] that, if one takes the risk-neutral criterion of maximizing expected assets at an horizon point T (and if overdrawing at the same interest rate is permitted), then one can optimize the

https://doi.org/10.1017/S1365100502027025 Published online by Cambridge University Press

enterprise in isolation, assuming discounted costs. However, if one takes a risk-sensitive criterion, then this is simply not the case; in running the enterprise, one has to consider the implied distribution of assets between enterprise and bank balance as a nontrivial allocation problem. That is, one has a certain source of gain (the bank account) and an uncertain source (the enterprise), and one is indeed then sensitive to the differing risks presented by these.

We can give a compressed version of the mathematics for those who want more detail. Suppose that operation of the enterprise can be formulated as a Markov decision process with state variable x and decision variable u. For simplicity, we assume it to be time homogeneous, with instantaneous cost function $c(x, u)$. Suppose that this operation is backed by cash in an interest-bearing account, in that if w_t is the amount in the account at time t, then

$$w_{t+1} = \beta^{-1}(w_t - c_t), \tag{15}$$

where c_t is the value of $c(x, u)$ at time t. The subtraction of costs (or deposit of gains, if c_t is negative) represents the cash flow between account and enterprise; the factor β^{-1} represents the growth of the account balance by compound interest.

Suppose that we regard the problem as a purely terminal one, in that the policy π is chosen to maximize $E_\pi[\mathbf{C}(x_T, w_T)]$, a function purely of x and w at the horizon point T. The problem then has a joint state variable (x, w), and decisions amount as much to investment decisions (i.e., allocation of assets between enterprise and cash account) as to the running of the enterprise itself. The dynamic programming equation for the joint system is

$$F_s(x, w) = \sup_u E\left(F_{s-1}\{x', \beta^{-1}[w - c(x, u)]\} \mid x, u\right), \tag{16}$$

where F is the value function, $s = T - t$ is time to go, and x, x' and u are the values of x_t, x_{t+1} and u_t.

Suppose that the terminal reward is simply the value of terminal assets,

$$F_0(x, w) = w - F_0(x), \tag{17}$$

where $-F_0(x)$ is the surrender value of the enterprise when it is in state x. An inductive argument then shows that the dynamic programming equation (16) has the solution

$$F_s(x, w) = \beta^{-s}[w - F_s(x)] \tag{18}$$

where $F_s(x)$ obeys the equation

$$F_s(x) = \inf_u\{c(x, u) + \beta E[F_{s-1}(x') \mid x, u]\}, \qquad (s > 0) \tag{19}$$

This is exactly the dynamic programming equation for the enterprise in isolation, but with discounted costs.

https://doi.org/10.1017/S1365100502027025 Published online by Cambridge University Press

However, if one were to choose some risk-sensitive measure of terminal utility (or even if overdrawing carried a penalty differing from the levy of interest at the same rate as it is earned on a positive balance), then one would not have this reduction. One would be left with the more general equation (16), indicating that allocation of assets between enterprise and account remains an issue.

6. IMPERFECT OBSERVATION, A SUBTLER CERTAINTY EQUIVALENCE

Assume that current state is not observable, but, rather, that at time t one observes

$$y_t = Cx_{t-1} + \eta_t \tag{20}$$

Here, η is observation noise; we shall suppose the elements of the sequence $\{\eta_t\}$ are independent normal with zero mean and covariance matrix M, supposed for simplicity, to be independent of $\{\varepsilon_t\}$. Then, expression (9) for the discrepancy must be modified to

$$\mathbf{D} = \frac{1}{2}(x_0 - \hat{x}_0)^T V_0^{-1}(x_0 - \hat{x}_0) + \frac{1}{2}\sum_{t=1}^{T}\left[\varepsilon_t^T N^{-1}\varepsilon_t + \eta_t^T M^{-1}\eta_t\right],$$

where the first term represents a starting assumption: At time $t = 0$ the value of x_0 is known only to be normally distributed with mean $\hat{x}_0$ and covariance matrix V_0. We can substitute $\varepsilon_t = x_t - Ax_{t-1} - Bu_{t-1}$ and $\eta_t = y_t - Cx_{t-1}$.

For the general case the point is simply that $\mathbf{C}$ is a quadratic function $\mathbf{C}(x, u)$ and $\mathbf{D}$ a quadratic function $\mathbf{D}(x, y, u)$, where x, y, and u denote the sequence of these variables from starting point to horizon point. In the risk-neutral case at a given time t, one forms estimates of all current unobservables by minimizing $\mathbf{D}$ with respect to them. One then determines the optimal value of u_t by minimizing $\mathbf{C}$ with respect to currently undetermined controls, unobservables being replaced by their estimates. The fact that this indeed leads to the optimal evaluation of u_t is the *certainty equivalence principle*. If state structure is assumed, then all these determinations can be given a simple recursive form—the Kalman filter for the estimate and the dynamic programming determination of the optimal deterministic control.

What might be seen as an alternative statement of the certainty equivalence principle is the *separation principle*: One determines optimal estimate and optimal deterministic control separately, and then couples them just by the substitution of estimates for unobservables in the control rule.

For some years, no equivalent calculation could be found for the risk-sensitive case. The solution was found by Whittle (1981). The first point established was that the optimal control at a given time was determined by the minimization/extremization of stress described earlier. The second point established was that, in the state-structured case, there is indeed a principle by which one can separate determination of estimate and control, in a sense, and compute these determinations recursively. One does so by, at time t, dividing stress into past and future components, and assuming provisionally that current state x_t is known. Then,

https://doi.org/10.1017/S1365100502027025 Published online by Cambridge University Press

minimization/extremization of past and future components can be carried out separately and recursively. The only unobservable that has not been extremized out is x_t itself, and one then couples the two calculations by doing just this, thus determining an effective estimate $\hat{x}_t$.

There is thus indeed a separation principle, and the coupling of the two calculations in the way described (or, rather, the fact that the whole procedure leads to an optimal determination of u_t) is the generalization of the concept of certainty equivalence that is required. As we saw in Sections 2 and 3, this leads to conclusions that are strange if one's intuition has been formed by the risk-neutral case; noise statistics now affect control and costs affect estimates. It could be said that certainty equivalence holds in the sense of *free extremization*: that determination of the optimal control u_t as a function $u(W_t)$ of current observables is now replaced by a free extremization of stress with respect to currently undetermined controls and current unobservables. See Whittle (1990a, 1991b) for extended treatments.

7. ROBUSTNESS

The initial interest of a risk-sensitive formulation is that it gives weight, not merely to the expected value of cost (or reward), but also to its variability. The development of the "phantom other" view enriches the picture. However, what is also interesting is the robustness (i.e., proofing against model misspecification) induced by a degree of risk averseness.

That there is such an effect is confirmed by the bridging with the H_∞ criterion, which by its nature looks for the worst-case scenario. An alternative point that emerges is that, if the model has been optimized under the assumption of imperfect state observation with observation error, then this in itself guarantees a degree of robustness—essentially because the response of the controlled model to observation error is much the same as its response to a slight model misspecification.

One might note also a special property of the H_∞ norm: that $\|GH\|_\infty \leq \|G\|_\infty \|H\|_\infty$. This property in itself makes analysis of the robustness question much easier.

The robustness implications of an H_∞ formulation were analyzed by Zames (1981); see the expositions by Francis (1987) and Doyle et al. (1992). The corresponding implications for a risk-sensitive formulation have been pursued by a number of authors; see, e.g., Boel et al. (1997), Malcolm et al. (1999).

8. THE ESCAPE FROM LQG ASSUMPTIONS—LARGE DEVIATIONS AND PATH INTEGRALS

It is natural to ask whether these ideas are totally dependent upon LQG assumptions or whether they hold more generally. In fact, there is a remarkably natural general version, following immediately from one of the most significant developments in pure and applied probability over the past 30 years—the theory of *large deviations*. [See, e.g., Bucklew (1990), Shwarz and Weiss (1995).]

https://doi.org/10.1017/S1365100502027025 Published online by Cambridge University Press

The equivalent of $\mathbf{D}$ for a general case will be a rather messy object. However, suppose that the system has a large *scale* n, and that a state variable such as x can be viewed as a system average. Then one can, in fact, evaluate a function $\mathbf{D}(x)$, known as the rate function, with the property that

$$P(x \in \mathcal{A}) = \exp\{-n[\inf_{x\in\mathcal{A}} \mathbf{D}(x)] + o(n)\}.$$

More than this, one has asymptotic evaluations of expectations of the type

$$E[\exp(-\kappa\theta\mathbf{C}(x)] = \exp\{-n \inf_{x} [\theta\mathbf{C}(x) + \mathbf{D}(x)] + o(n)\}.$$

This is then perfectly set up for an risk-sensitive treatment and all ideas go through in this natural asymptotic sense. These ideas are followed through by Whittle (1990b, 1991a, 1996b). The analysis of these papers has since been rigorized or extended in a number of directions by several authors; see, e.g., Fleming and Soner (1992), James (1992), James et al. (1994), and Fleming and McEneaney (1995).

Brute-force application of a stress-extremizing principle in cases where it is not very well supported theoretically still gives interesting results [Whittle (1994)]. Consider, for example, the simplest classical version of the optimization of consumption. We phrase this in continuous time, so that assets x follow the equation

$$\frac{dx}{dt} = ax - u,$$

where a is the continuous-time interest rate and u is the rate of consumption. We shall suppose the utility function

$$\mathbf{U} = -\mathbf{C} = \int_0^T \log(u)\, dt + \kappa \log[x(T)]$$

for a fixed termination time T. Then, one finds by classical methods that the optimal rate of consumption is related to current assets x by

$$u = \frac{x}{s + \kappa}.$$

Here $s = T - t$ is time-to-go, that is, residual lifetime. However, none of us knows the hour of our death. Suppose that we must base decisions on current assets x and an *apparent residual lifetime* y. Here, we suppose that $dy/dt = -1 + \varepsilon$, and that death takes place when y first equals zero. Then we find that a slight degree of optimism has the effect on consumption illustrated in Figure 2.

For most values of x and y the effect is counterintuitive; optimism makes one consume less. The reason is that optimism leads one to suppose that one will live longer, thus encouraging saving. For small assets x, there is a band of y for which it is optimal to *increase* consumption. This is when apparent residual lifetime is

https://doi.org/10.1017/S1365100502027025 Published online by Cambridge University Press

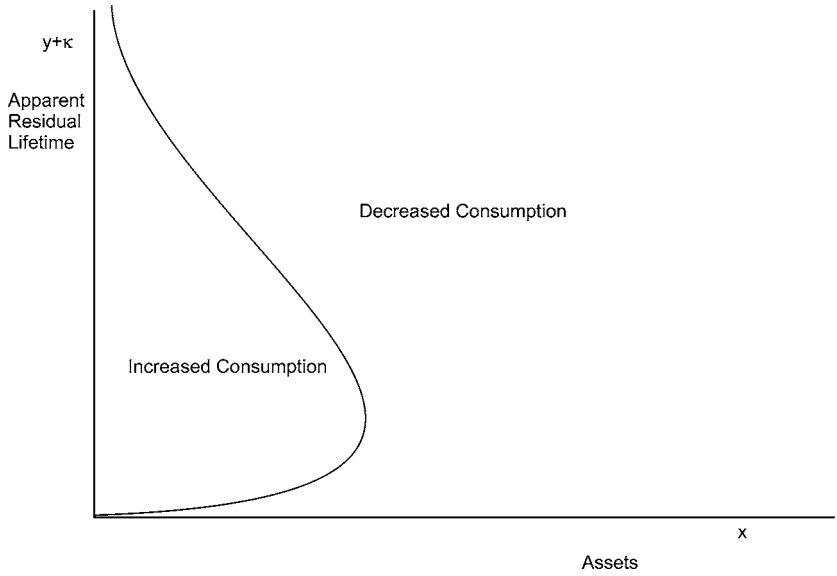


FIGURE 2. Effect of an increase in optimism on the optimal consumption rate.

neither so large that there is advantage in trying to build up equity nor so small that assets are more than sufficient for the time remaining.

REFERENCES

Bertsekas, D. (1987) *Dynamic Programming: Deterministic and Stochastic Models*. Prentice-Hall.

Boel, R.K., M.R. James & I.R. Petersen (1997) Robustness and risk-sensitive filtering. In *Proceedings of the 36th IEEE Conference on Decision and Control*.

Bucklew, J.A. (1990) *Large Deviation Techniques in Decision, Simulation and Estimation*. New York: Wiley.

Doyle, J.C., B.A. Francis & A.R. Tannenbaum (1992) *Feedback Control Theory*. New York: Macmillan.

Fleming, W.H. & W.M. McEneaney (1995) Risk sensitive control on an infinite time horizon. *SIAM Journal of Optimal Control*. 33, 1881–1915.

Fleming, W.H. & H.M. Soner (1992) *Controlled Markov Processes and Viscosity Solutions*. Berlin: Springer-Verlag.

Francis, B.A. (1987) *A Course in H_∞ Control Theory.* Berlin: Springer.

Glover, K. & J.C. Doyle (1988) State-space formulae for all stabilizing controllers that satisfy an H_∞-norm bound and relations to risk sensitivity. *System & Control Letters* 11, 167–172.

Jacobson, D.H. (1973) Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games. *IEEE Transactions on Automatic Control* AC-18, 124–131.

Jacobson, D.H. (1977) *Extensions of Linear-Quadratic Control, Optimization and Matrix Theory*. New York: Academic Press.

James, M.R. (1992) Asymptotic analysis of nonlinear stochastic risk-sensitive control and differential games. *Mathematics of Control Signals Systems* 5, 401–417.

James, M.R., J.S. Baras & R.J. Elliott (1994) Risk-sensitive control and dynamic games for partially observed discrete-time nonlinear systems. *IEEE Transactions on Automatic Control* AC-39, 780–791.

https://doi.org/10.1017/S1365100502027025 Published online by Cambridge University Press

Malcolm, W.P., R.J. Elliott & M.R. James (1999) Risk sensitive filtering with continuous time observations. In *IEEE Conference on Decision and Control*, pp. 143–150. Piscataway, NJ: IEEE Press.

Shwarz, A. & A. Weiss (1995) *Large Deviations for Performance Analysis*. London: Chapman and Hall.

Whittle, P. (1981) Risk-sensitive linear/quadratic/Gaussian control. *Advances in Applied Probability* 13, 764–777.

Whittle, P. (1990a) *Risk Sensitive Optimal Control*. Chichester, England, and New York: Wiley.

Whittle, P. (1990b) A risk-sensitive maximum principle. *Systems and Control Letters* 15, 183–192.

Whittle, P. (1991a) A risk-sensitive maximum principle: The case of imperfect state observation. *IEEE Transactions on Automatic Control* AC-36, 793–801.

Whittle, P. (1991b) Likelihood and cost as path integrals. *Journal of the Royal Statistical Society* B 53, 505–529.

Whittle, P. (1994) Risk sensitivity, large deviations and stochastic control. *European Journal of Operational Research* 73, 295–303.

Whittle, P. (1996a) Why discount? The rationale of discounting in optimization problems. In C.C. Heyde (ed.), *Athens Conference on Applied Probability and Tame Series*, *Lecture Notes in Statistics*, Vol. 114, pp. 354–260. Berlin: Springer-Verlag.

Whittle, P. (1996b) *Optimal Control: Basics and Beyond*. Chichester: Wiley.

Zames, G. (1981) Feedback and optimal sensitivity. *IEEE Transactions on Automatic Control* AC-26, 301–320.

https://doi.org/10.1017/S1365100502027025 Published online by Cambridge University Press