Optimal Feedback Control Rules Sensitive to Controlled Endogenous Risk-Aversion

Dan Protopopescu[1 2]
Center for the Study of Organizations and Decisions in Economics
Universitat Autònoma de Barcelona
E-mail: *dan.protopopescu@uab.cat*

June 2008

Abstract

The objective of this paper is to correct and improve the results obtained by Van der Ploeg (1984a, 1984b) and utilized in the theoretical literature related to feedback stochastic optimal control sensitive to constant exogenous risk-aversion (see, Jacobson, 1973, Karp, 1987 and Whittle, 1981, 1989, 1990, among others) or to the classic context of risk-neutral decision-makers (see, Chow, 1973, 1976a, 1976b, 1977, 1978, 1981, 1993). More realistic and attractive, this new approach is placed in the context of a time-varying endogenous risk-aversion which is under the control of the decision-maker. It has strong qualitative implications on the agent's optimal policy during the entire planning horizon.

JEL Classification: C51, C61, C91, D81.

Keywords: Controlled stochastic environment, rational decision-maker, adaptive control, optimal path, feedback optimal strategy, endogenous risk-aversion, dynamic active learning.

1. Introduction and Survey

Behavior optimization is very important in any positive economic analysis. The need to solve optimization problems arises in almost all fields of economic inquiry. In the real world, it exists many situations in which a rational decision-maker is constrained to face up to multiple uncertainties relative to the behavior of a dynamic stochastic environment. In general, the decision model either makes assumptions on how the decision-maker will reponds to his environment or derive this behavior from optimality-based considerations.

Uncertainty is an intimate dimension of economics. This generally arises because of inherent difficulties of perception and information processing. We are generally uncertain about the structure of the model, the numerical values of its parameters of interest and the future values of exogenous or random variables.

It is well-known the crucial role that the information plays in the decision making process of individual agents facing uncertainty. More information is always beneficial. It cannot have a negative value. Incorporation and judicious use of further prior information into the statistical procedures will produce better estimators. Greater information will, in general, reduce the environmental complexity and hence the decision-maker's uncertainty.

In practice, the decision-maker bases his decisions on some body of knowledge. He does not know which state in the future will in fact hold. In this sense, the knowledge must be viewed as

[1]Submitted for publication to *Journal of Economic Dynamics and Control*.
[2]Financial support from the Spanish Ministry of Education and Science through grant PTA-2003-02-00877 and FEDER is acknowledged.

future oriented expertise. When that base of knowledge evolves over time, regulatory decisions evolve over time. In dynamic behavior situations, generally the uncertainty is only gradually solved through time and the decision-maker may be afforded continuously the opportunity to revise his plan of action. Under uncertainty, the optimal sequence of decisions depends on not only the expected losses, but also the flexibility, in terms of availability of future options associated with each decision. Does not matter the preferred decision-maker's policy, the uncertainty will alter this considerably. We recall here some efficient methods in order to reduce the uncertainty, as the control of the future, the increased power of prediction, or by diffusion (see, KNIGHT, 1971).

The theory has generally little to say about where and how the error terms propagate in equations. Roughly speaking, the decision-maker face up to diversifiable errors from the system. They can arise from several sources. Real world is never free of disturbance.

In practice, the true distribution of the data (or a parametric family containing it) is never precisely known. Uncertain disturbances and parametric variations cannot be specified ahead of time. In other words, the decision-maker does not know ex-ante how these will be resolved. The evolution of the dynamic stochastic process is partially subject to the control of the economic agent who is imperfectly informed as regards the law of motion. The information acquired is perturbed by unobserved random shocks with a more-or-less known distribution. As regards the parameters of interest, they are not individually observable and hence they are not known with certainty.

For econometric purposes, it is often assumed that a steady state is always observed for the environment state equation. In other words, the parameters of interest are supposed temporarily unstable. However, in the absence of stationarity, the dynamics become much more interesting. Until the work of ROSENBERG (1968), a little attention was given to the fact that the parameters in the econometric relationships are likely to vary systematically over time. He explored a number of ways in which misspecification may lead to parameters variation over time. As long as the dynamic effects of parameters variations are slow in comparison to state variations, control design can be based on a time-invariant dynamic model. Fast parameters may be indistiguishable from state components, in which case the parameters should be included in an augmented state vector for estimation and possibly for control.

The bayesian analysis can be viewed as a potentially attractive approach for this type of modelling. The analysis of economic policy problems from a decision theoretic point-of-view is more satisfactory in this context. The point estimates are replaced by the entire posterior distribution. In the real world, the priors of the decision-maker may change. Incorrect beliefs about the unknown parameters will induce a significant bias in the controls and target variables. For a bayesian decision-maker, the uncertainty has several dimensions. It may manifest in incomplete knowledge about the response parameters, in incomplete learning from sample estimates of parameters from a finite set of observations, or due to the model stochasticity.

However, to base decisions on estimates of the parameters that characterize the decision problem is not always optimal. We can find, in this sense, numerical examples which can prove that basing decisions on conventional parameter estimates can leed to large losses of decision-maker's utility (see, KLEIN ET AL., 1978).

Even if the agent "knows" the parameters of the model, it is possible that he cannot observe perfectly the stochastic process (e.g., the state variable is observed with noise), or that the information is incomplete (e.g., he has an incomplete signal about the current state). As a consequence, the forecasts may be dispersed. Moreover, supposing that the agent can observe "perfectly" the environment, the true model generally depends on other exogenous variables which escape to the control. These can be observed completely or not, the agent can be conscious or not of their existence. Missing data will have a negative impact on the estimated parameters or statistics and thus, on the optimality of the control instruments. There will hence be a considerable deviation of the system from its optimal reference path. On the other hand, the process of learning will be much more slower due to the loss of information.

The decision-maker can utilize the history of the process and develop an approximate model in order to analyze the behavior of the system. Unless we profit of the particular structure of the problem, it will generally be very expansive to generate information to approximate a (large) model with enough accuracy. To include this as part of an iteration procedure would be an order of magnitude more complex. Satisfactory approximations are difficult to obtain. A close approximation is sufficient but not necessary for solving the problem. There is a trade-off between the utility of a better approximation and the increase in computational costs which limit our ability to study the model in a data relevant manner. Generally, we have a hierarchy of costs for different levels of approximation of the "true" model. The agent seeks to model the main features of the data generating process in a relatively simplified representation (which becomes gradually more complicated if additional data become available) based on the observables and related to prior economic theory. The problem is whether this simplification does or does not involve a loss of information.

Because of the measurement error and other random effects, there is considerable uncertainty in determining whether the observed data actually are generated from the true model. Economic models are only rough approximations of the true data generating process, generally unknown. It is useful to note here that a local linear approximation of the model is made around a random point (due to the error term) near of the true value which is not known before the decision period. Local here means that only the behavior of the model in small neighborhoods of suitably chosen points is considered. For a narrow neighborhood of the optimum point, the linearity hypothesis is more likely to be a good approximation. If the model is without stationary state, we can iteratively determine the point about which to approximate. The question is how well this approximation might work. We point out that the learning aspect demandes a study on the robustness of the strategies obtained by linearization. In this sense, it exists numerical algorithms which permit the respecification of the criterion after each maximization (see, Rustem, 1979). It will increase the precision of parameter estimates. How useful are the models based on approximate solutions to optimal behavior, this is a question whose response is given by the various numerical applications.

The mathematical theory of stochastic optimal control is well developed (see, Fleming & Rishel, 1975, Kendrick, 1981 and Karatzas & Shreve, 1988, among others). Mathematically speaking, an optimal control problem is concerned with the determination of the best ways to achieve a set of objectives as indexed by a criterion function when the performance is judged over many periods and when the dynamic behavior of the system is subject to a set of constraints. In the terminology of the control theory, the variables are divided into those which represent the condition or state of the objective functional (the so called state variables), and those which guide or control the state variables (the so called control variables).

Depending on the difficulty of the problem, the dynamic optima can be either theoretically analyzed or empirically tested. Improvements in finding closed-form solutions of dynamic stochastic models is still very slow. In order to obtain analytically tractable results, restrictions which are less attractive from an economic point of view have to be imposed.

The only model which can be solved in any generality is the linear-quadratic approximation model which gives linear decision rules under given specific conditions, very convenient for theoretical analyses and attractive on computational grounds.

Linear models are widely used in the literature due to their theoretical simplicity and flexibility or for avoiding hard numerical estimation. One can think of a dynamic model as being linear if its global properties can be completely characterized by its local behavior. Non-linear dynamic models do not have this property of equivalence between local and global dynamics and thus are substantially more complicated to analyze. The non-linearity typically impedes analytical solutions for an optimization problem. The non-linear modelling is generally less amenable, especially by the presence of uncertainty. This is an other reason for which the literature is generally focused on the linear model.

Non-linearity may arise in diverse ways in the econometric applications and there are many possible approaches for specifying non-linear models. We can have, for example, non-linearity in parameters and /or in variables (see, AMEMIYA, 1974) as well as non-linearities in time series or with respect to the system disturbances.

In non-linear dynamic models (described by non-smooth functional forms), very special assumptions have to be made in order to obtain closed-form solutions. If the dimension of non-linear models is high, then the analytical treatment becomes very difficult. Moreover, they present considerable difficulties in terms of initialization and convergence. The existence of the optimum as well as the speed of convergence of the algorithm is restricted to certain configurations of the initial parameters of interest. It implies an adjustment mechanism of the tatonnement type.

The deviation caused by non-linearities in the model is quite important in the sense that the stochastic optimal trajectory does not follow the desired path closely. The more non-linear the model is, the more difficult will be to track the targets. As flexible as the non-linear dynamic model may be, there is a substantial specification uncertainty.

When the analytical formula cannot be obtained easily, the analysis of the problem requires the use of some numerical computational algorithms or simulation techniques-based methods. They remain the only viable way to obtain insights about the system studied. However, difficulties in numerical computation arise, because the dynamic optimization problem may be characterized by multiple optima. It generates difficulties as regards the numerical implementation.

Confrontation with data is very important. The theoretical model must be consistent with the empirical model. In recent years, a number of numerical methods have been proposed in the literature of stochastic simulation (see, TAYLOR and UHLING, 1990, MARCET, 1994, AMMAN, 1995, RUST, 1996 and JUDD, 1998, among others). The continue increase in the computers computational speed makes feasible new adaptive control rules /learning algorithms (designed for experimentation) and enlarges the class of the models that can be approached by simulation using the data generating process. They are playing an increasingly important role in economic analysis (especially in controlling economic dynamics) and allow to gain experience from large structural models (whose properties are largely revealed by empirical experimentation). In order for numerical experiments to be effected rigorously, it is important to dispose of error bounds or accuracy estimates of the computed solutions (see, MANUEL and AGUIAR, 1998).

The dynamic programming method (in discretized version) provides in this sense, a constructive recursive procedure for computing the optimal decision rules. This procedure (based on a process of backward induction) amounts to a solution algorithm that allows us to obtain numerical solutions to specific problems as well as analytic characterizations of a wide class of problems.

Unfortunately, the amount of computation required to obtain the dynamic programming solution rises exponentially with the number of variables in the model, due to "the curse of dimensionality" (see, PITCHFORD, 1977). The handling of policy instruments requires information on all states of the system, making the policy rule complicated from the point of view of implementation. It is more difficult to investigate the properties of the optimal policy when one allows for complex history dependence.

It is therefore essential to maintain a balance between the desire for a more sophisticated economic model and the need for nominal configuration in terms of computer use. This is what makes the research in this area both difficult and interesting. In order to control the simulation of the optimal policy, it is very important to use the analytic information from asymptotic theory.

A weakness of these simulation techniques-based methods is that the properties of the model may depend on a particular specification of the true model (i.e., a particular set of parameters and functions chosen for simulation), and one may get a distorted picture of the properties. In general, when information is gathered in the real world, the data generating process is not independent and free of noise.

The simulation results must not be considered as a perfect substitute of the theory but only as an instrument which may confirm the theoretical results. If an anomaly arrives, this is presumably due to a particular drawing of pseudo-random numbers (it exists uncertainty in the choice of experiments). It is possible that two decision-makers legitimately make substantially different inference from the same data. In general, the transition from theoretical models to empirical models is severely constrained by the quantity and the quality of the available data. The data may not contain enough information to decide a particular issue. The bias of the sample selection is viewed as an error of specification (see, HECKMAN, 1979) and a source of numerical instability in the empirical model, leading to loss of significant digits in some or all results.

In the most practical situations (typically in parametric models), a less complicate model is likely to be preferred if we wish to pursue the accuracy of the estimation or to profit of important analytic advantage. There is, for example, a gain of information from the theoretical analysis of the linear model viewed as an a priori specification. It may serve as a good illustrative theoretical example by simplifying the analysis considerably. Thus, in the context of a game, the linearity of the model allows us for a complete characterization of the set of equilibria (which is not the case for a non-linear model). The assumption of linearity in functional relationships serves to simplify the conceptual and computational development of the theory.

Efficient dynamic specification tests seem to be relevant for the model selection. In principle, any dynamic optimization model is empirically testable. It allows to study the behavior of the model under different environments. Empirically, there are often conflicts in the criteria of the selecting a model to achieve multiple objectives. Several procedures exist for testing the specification of an econometric model in the presence of one or more other models which purport to explain the same phenomenon (see, DAVIDSON and MACKINNON, 1981, among others).

It is useful to note that the rejection of the null hypothesis should not lead to automatic acceptance of the alternative hypothesis, as the test could have a greater power against other deficiencies (see, SARGAN, 1988). In other words, the fact that the test fails to reject an hypothesis should not necessarily leads to accept it. Thus, the linearity may not be rejected (when testing for linearity) if, for example, other variables are added to the initially specified linear model (see, GRANGER and TERASVIRTA, 1993). By consequence, it is helpful to know when we can decide that the non-linearity is the element which causes misspecification in the linear model.

The linear approach is not preferable when it works with a naive and extremely simplified model (too simple to capture the effects of environmental uncertainty). It is necessary to have a robust specification of the linear model. It is also well-known that misspecified theoretical models could forecast well if the process remains constant, while good models could forecast poorly if the data variance is high. In other words, a model can be acceptable despite having a poor fit and, the fact that a rival model has a better fit does not necessarily make it a better one (see, HENDRY, 1995). Sometimes, it happens that although the model is restrictive and in some ways unrealistic, it brings out many of the key insights. If a model is found to be superior, the matter which remains to be solved is to prove if the difference between the two specifications is significant. Naturally, all the results of the linear model must be asymptotically valid in the non-linear specification case. This is the encompassing principle which requires a model to be able to explain characteristics of rival models. However, a model including another does not necessarily encompass this model (see, Gourieroux and Monfort, 1995).

When parameterizations are not a part of prior knowledge or the space of parameters has high dimension, a nonparametric approach is more appropriate. It allows for a more robust model specification. The methods based on nonparametric smoothing provide valid inference under a much broader class of structure letting speak the data without constraints and dictate from themselves the true form of the regression instead to specify the functional form of this (or the distribution of the residuals). The regression function is not so reduced to a class of functions and preserves all its freedom. However, in the case of a non-parametric decision problem the agent (in fact, the modeler) has less information about the uncertainty distribution than in the parametric problem. The higher the set of parameters, the less informed the agent. Moreover, the robustness of nonparametric methods is not free and nonparametric statistics are less efficient than traditional parametric methods.

Sometimes, even a nonparametric model may not be identified without imposing some restrictions and thus making it semiparametric. In this case, it is useful to use a prior information on the regression shape in the form of a parametric model (the functional of the interset is smoother) in the nonparametric regression, that is, we nonparametrically encompass the parametric model (see, Gozalo and Linton, 1995). A semiparametric model involves a finite dimension parameter and an unknown nuisance function. This kind of models combines the flexibility of a nonparametric model with the efficiency property of a parametric model.

Each econometric approach has its advantages /strenghts and limitations /weaknesses. New approaches rise new difficulties. The objective of the modeller is to discover the most appropriate model that explains the observed data. In general, there are many implicit restrictions derived from the economic theory. Any specification would be preferred, the builded model must be consistent with the economic theory, data admissible, congruent with the data and computationally attractive (see, Hendry, 1995).

Even if the decision-maker's strategy is perfect, the control can still be improved because we cannot completely eliminate the uncertainty from the system. Various difficulties are encountered by the decision-maker when modelling real economic phenomena. There are multiple sources of uncertainty that he must deal with by using optimal adequate solutions.

We conclude from this short methodological introduction that there is no royal way to develop good models. In other words, there are no precise rules for econometric model design. The economic theory is generally based on restrictive hypotheses, non-verified in totality. Usually, they are technical hypotheses, convenient for theoretical and analytical purposes. Simple hypotheses are very rare in econometrics and these reflects the limits of the modelling. In general, stronger results require stronger assumptions and thereby harder informational requirements. Very often in the literature, arbitrary and untested hypotheses are chosen by the modeller for their practical convenience. The more general approach is one with a set of minimal and feasible assumptions but usually various difficulties are encountered in this case. The modelling is only an additional tool of observation, and it cannot describe an economic phenomenon in its globality. Statistical and economic models are only rough approximations of the true data generating process. Divergences often arise between theory and experimental evidence. Theoretical predictions are sometimes compared to empirical results although they do not fulfil the same set of assumptions. This is mainly due to the absence of an unified theory. Complex real world interactions between economy and environment is the main barrier to applied research within the field of economic modelling.

This paper is organized as follows. Section 2 deals with the problem statement and makes preliminary considerations. Section 3 presents the model. Section 4 deals with the probabilistic hypotheses on the acquisition of information. Section 5 corrects the theoretical results of Van der Ploeg (1984a, 1984b) on the estimation of the feedback optimal policy in the context of a constant exogenous risk-aversion. Section 6 introduces the concept of endogenous risk-aversion. Section 7 improves the formulas obtained for the optimal feedback control rules in Section 5, by considering the more realistic case of a time-varying endogenous risk-aversion subjected to the control of the decision-maker. Section 8 draws some conclusions and makes suggestions for further research.

2. Problem Statement and Preliminary Considerations

Facing a risky environment, a rational decision-maker disposes of an optimal set of control instruments in order to constrain the system to follow a fixed optimal trajectory which ensures its equilibrium and stability. The goal is the path. There will generally exist a trade-off between the efficiency of the control instruments and the decision-maker's fixed objectives in an uncertain an changing world.

The rationality of the decision-maker is characterized by the anticipation that the environment will be affected by other factors than the control instruments. It implies an forward-looking behavior. These factors are completely or partially observed and may be exogenous or endogenous variables as well as continuous unobserved random shocks. This can also be viewed as consistent choice /action or the pursuit of the decision-maker's self-interest (see, Walsh, 1996).

In a more general context, the instrument can be used for experimenting, the goal being to learn the "true" parameter of interest. Strategic experimentation is an important aspect of optimal decision making for a wide class of learning problems. The purpose of the experimentation is to gain additional information (which is valuable for future decisions) in order to obtain an optimal learning. Optimal control with learning about unknown parameters has been applied to a variety of economic problems (e.g., optimal investment with production uncertainty, monopolistic pricing with unknown demand, fiscal and monetary policy with imperfect knowledge about the macroeconomy). If the cost of information is too expansive for permitting the learning (e.g., the model is high non-linear and we search for the better linear approximation, or the system uncertainties are large), we can be pleased with a rational random behavior (see, Barbosa, 1975). If the dynamic environment is highly sensitive to non-rational actions, then the stochastic control will be optimal if we can reconcile the desire of the risk with the non-stationarity of the process and the instability of the equilibrium. Optimal decisions generally involve a certain degree of experimentation. The more the decision-maker cares about future performance of the dynamic process, the more he will experiment. The objective of the decision-maker, in this case, is to determine the optimal level of policy experimentation. A rapid decline in the variability of the system state can be associated with an optimal experimentation. This substantially improves the speed of learning and the bias in the control and target variables (see, Wieland, 2000).

In practice, decisions are based on parameters which are not known with certainty and may vary over time. When parameter uncertainty is large, experimentation becomes significantly important. It increases with the variance of the unknown parameters. The degree of experimentation is expected to be smaller with time-varying parameters than with constant ones. In contrast to the constant fixed parameters case (when the incentive to experiment is temporary; it disappears over time as parameter estimates become more precise), the incentive to experiment remains high and never ceases when parameters vary over time (see, Beck &

Wieland, 2002). On the other hand, doesn't matter the type of specification, the incentive to experiment will naturally increase with the variance of the random shocks as well as their degree of persistence.

The notion of rational decision making in an uncertain environment is associated with the expected utility-function maximization behavior. Rationality lies in the correspondence of the decision-maker's action with some goal or objective. For example, he does not refuse to act in accordance with the efficient outcome (that is, at best of his interest). However, the decision-maker's preferences are generally incomplete. It is very rare in econometrics to be able to fully specify the utility function. No decision-maker has sufficient a priori knowledge to fully specify his utility function.

The decision model is based on the joint use of the econometric model and of the decision-maker's preference function. The latter will be optimized under the constraint represented by the former and, very likely, other necessary constraints. The decision will be not separated from the decision procedure and the judgment of rationality carries on the whole. The control rule will be characterized by informational requirements and the decision criterion. As regards the decision-maker's strategy, this is based on an adaptive expectation mechanism and on a feedback rule. Because the environment is generally non-stationary, it will influence the outcome of any decision, so that, the decision-maker will have to analyze the evolution of a dynamic system in which the present state is a consequence of the decision taken yesterday. His optimal actions are conditional on past history.

A decision problem under uncertainty consists essentially in establishing a preference ordering over a set of stochastic variables.

From dynamic economic theory we know that optimal decision rules vary systematically with exogenous changes in the structure of series relevant to the decision-maker. It follows that changes in policy will systematically alter the structure of series being forecasted by the decision-maker, and therefore, the behavioral econometric relationships as well. Important cumulative effects of the parameters change on the time path of the state and control variables will be present.

At each control period, the level of uncertainty of the decision-maker is given by the deviation of the actual state of the system from his local objective. High deviations from the fixed targets correspond to a high level of uncertainty. The decision-maker adjusts to keep small the difference between actual and assumed system characteristics by monitoring the system fluctuations. He optimally chooses the control instruments on the basis of a non-decreasing endogenous information set.

The optimality of the strategy is generally defined relative to the information the decision-maker has at the time the strategy is used. It exists so a relationship between the instruments efficiency and the optimal policy chosen by the agent. The decision-maker can use a knowledge base of past and present information to effect a control strategy, but future information is unavailable. Because exogenous shocks in the future are not predictable, the decision-maker strategy cannot incorporate them into the decision.

For the purpose of this analysis, it is assumed that the decision-maker emploies a closed-loop strategy which generally depends on the history of the process and thus includes feedback information. The decision-maker will constantly monitor the output of the process under control. The information is employed in real time. The knowledge upon which the decisions are based increases gradually with the passage of the time and due to the wisdom derived from experience. The decisions made in the past will be reflected in changes in the state of the system itself and they will influence the perception of the future actions to be analyzed. Because the source of randomness may differ from one application to another, the decision making response may vary.

It may be the case when the relevant information acquisition cost is high, most likely due to the permanent random shocks in the system or because of the slow inertia of the economic environment (generally, the environment changes very slowly in relation to the speed at which the economic agents learn). In a closed-loop strategy, the policy does not require some large periods of engagement from the part of the decision-maker. In other words, the control rules will be sensitive to the choice of the working horizon (a fixed number of stages in which the observations may be made, the agent's actions are taken, and the environment signals may be sent). Generally, the length of the working horizon does not only depend on the number of periods but also on the unity of measure chosen. A question remains: What is the optimal length of the planning horizon on which the decision-maker bases his decisions?

An infinite-horizon problem is not generally compatible with a closed-loop strategy. The conceptual and mathematical elegance of infinite horizon models is impractical for a computational viewpoint (even if the policy is easier to implement). To solve such a problem, it is initially convenient to consent ourselves with finite horizon approximates by including some terminal criterion. The advantage of a finite horizon also lies in the possibility to use forward recursive filtering techniques (see, Kalman, 1960, Kalman & Bucy, 1971, Anderson & Moore, 1979, and Harvey, 1990, among others) which allows to monitor the expectation formation process and implicitly the evolution of the stochastic system. In other words, the noisy observations are filtered in order to estimate the distribution of the state variable. This is specific to an incremental learning model based on a sequential forecast which varies with time and history. Adaptive filtering is one of the major contributions in modern stochastic control theory. It can be applied to treat unobserved components and data revision, or to generate innovation sequences. However, it must be viewed as a complement to econometrics methods, rather than as a substitute. It does not allow to manage large amounts of finite information. Its major drawback is that it generally restricts attention to the normal distribution.

It is useful to note that the solution found in finite horizon must converge to the solution found in infinite horizon. The infinite-horizon problem can be viewed as an approximation of the finite-horizon problem with a large planning horizon.

The decision-maker tries to reduce the endogenous uncertainty associated with his actions by acquiring information from the begining of the control to the moment of decision. He has the possibility to learn from errors and to make a self-evaluation of his actions. We can say that a closed-loop strategy is robust, in the sense that it anticipates the possibility of a disturbance. It can thus prevent unexpected shocks. The feedback control responds not only to the effects of the random inputs, but also to the measurement errors as well. Thus, it is not necessary to be able to identify and measure the sources of disturbance. If the decision-maker is interested in determining the effect of parameter changes on the optimal control policy, then the closed-loop strategy is generally the best way to do so. For optimal policy experiments and associated hypotheses testing of the optimal control problem, the closed-loop (or feedback) solution is also preferable. This type of strategy has the advantage to continuously improve the decision-maker's optimal policy. Also, for optimal policy experiments and associated hypotheses testing of the optimal control problem, the closed-loop solution will be preferable. We can see the implementation of the optimal closed-loop control as a form of integral control. Due to the imperfect information about the reaction of the system over time, it is perfectly reasonable to consider a maximization of short-term for the utility in the context of a closed-loop strategy.

The performance of the closed-loop control is superior to any open-loop control in the stochastic dynamic context (see, Cruz, 1975, Xepapadeas, 1992, and Wiedmer et al., 1996). More the period of control is longer, more the effect of cumulated errors on the agent's optimal policy is significant. If the system evolution is perturbed at each step by a random shock, the open-loop policy will not integrate this stochastic characteristic for computing the future decisions. The information purchased during the period of control (that is, the history of the process) is not taken into account by the agent so that he will lose the strategic learning. This will affect the optimal policy efficiency as it is adopted on a long-term. Only the information purchased before starting the control process will be utilized. All the errors on the initial state of the system will be intactly transmitted until the end of the control period. An other weakness of this policy can appear when the exogenous variables of the model are affected by coefficients which vary in time. The dependence of the coefficients realizations at each point in the horizon is not required. In general, any dynamic discrete-time problem may be reformulated as a static problem (only one-period time horizon) and the open-loop control solution may be derived.

The closed-loop strategy is a refinement of the open-loop concept. The open-loop and closed-loop strategies are equivalent only under the perfect forecast assumption, which is unrealistic, in most circumstances. In general, the closed-loop solution deviates from the open-loop solution. Disadvantages of the open-loop controls are that they require much information about the future development of the system and that they are not robust. It is by using a closed-loop strategy that economic theory can be exploited at best.

Dynamic feedback entails measurements, and these may be uncertain or indirect. With uncertain or indirect measurements, it is necessary to estimate the state history that is most likely to have caused the measurements. Thus, the control principles and the estimation principles can be used together to solve the stochastic optimal control problem. In other words, the control and estimation strategies can be designed concurrently (that is, one depends upon the other).

Learning is one of the three aspects of the agent's uncertainty problem (beside the parametric uncertainty and stochasticity) and has many dimensions. It can take place at various levels of a decision problem. As learning constitutes a form of economic estimation, it is desirable to develop learning algorithms in a context that allows for dynamic structure. Learning possibility can occur only in dynamic models and appears more likely with longer planning horizons. The relative efficiency of the learning generally depends on the method chosen. An optimal behavior may arise from a learning process. However, if the model is very noisy, then the potential for learning is limited. In function of the success of model approximation, the learning may be more or less efficient.

A double learning dynamic is taken into account in our analysis: one which describes how the decision-maker adjusts his behavior towards risk over time, and an other which reveals the impact of his optimal actions on the system performances. It is useful to note that the reinforcement or stimulus-response learning is not based in our modelling on the principle that actions which have led to good outcomes in the past are more likely to be repeated in the future.

Economic models involving learning often have the potential for converting independent shocks into correlated movements in observables. Models with learning induce persistent effects of transitory shocks. This is an important feature of models with stochastic endogenous fluctuations.

Accurate representations of the reality generally involve active learning, allowing economic agent to experiment. This is the case when decisions are made as much to acquire information. The timing of information is a crucial aspect. The agent can acquire additional information by receiving a noisy signal about the true state of the world. The degree of information embedded in the observation of the state variable generally depends on the values of the control variables, so that the extent of learning about the latent parameters can be influenced directly by the agent. He has some influence over the rate at which information arrives, so that his behavior may generate information. It can reduce the uncertainty with which he is faced. We can affirm that the active learning makes the agent more experienced over time. Learning from experience is a form of active learning (see, BALVERS and COSIMANO, 1990). In general, the uncertainty will depreciate the economic agent's activity and will produces a temporary stability followed by a longer or shorter period of adaptation in instability which implies for the agent an additional effort allocated in the active learning.

It is important to underline the benefits from active learning in a stochastic optimization model (see, EASLEY & KIEFER, 1988, or KIEFER & NYARKO, 1989, among others). We will incorporate the value of learning into the optimization problem. The algorithm will anticipate future learning when choosing the control for each period, and thus will perturb the system early in time in order to reduce the variance of the parameters estimated later in time.

The control process is limited by the speed with which the decision-maker reacts to cautious changes in the environment. We speak here about the inertia of the decision-maker to non-significant environmental changes. This is concretized in deleted observations of the decision-maker. He generally reacts to sudden shifts of the dynamic system. On the other hand, there is an inherent inertia effect of the environment due to its capacity of reaction.

Control actions adapt as a consequence of changes in endogenous variables and also affect the observability of the system. It will exist, by consequence, feedback between the decision-maker's instruments and the system target variable. It will implicitly generate a short-run causality chain. When the variables are not perfectly related, or they are not perfectly controlled, there are more possibilities for observing causality. Sometimes apparent causality occurs because of the presence of unobserved variables. In any discussion of causality, the timing of when things happen is of crucial importance. It must put variables when they occur rather than they are first observed.

The decision-maker's optimal control has a dual effect. It reveals information as regards the state and parameters of the system while achieving the optimal objective. It is important to note that the decision-maker's actions are taken in real time, whereas his decisions will usually be formulated in advance. In other words, the time passes between taking a decision (which generally implies a selection from several known alternatives) and its implementation. This selection implies the existence of an well-defined goal.

3. The Model

Consider a stochastic data generating process managed by a system of discrete dynamic simultaneous equations. Several theoretical and empirical arguments can be enumerated in support of this model specification:

i) The current practice in the specification of parametric econometric models places considerable emphasis on the simultaneous relationships between variables. The economy is a system and the simultaneity is a property of the system. This is the most obvious reason for joint modelling and was the basis for the advances in econometrics started firstly by HAAVELMO (1944).

ii) Due to the sheer complexity of the behavior of an economic system, we very often need a dynamic specification (that is, there are lagged dependent variables among the regressors in the model). The use of distributed lags in econometric research is quite old (see, KOYCK, 1954). An advantage of this type of specification is that the time lags of a variable can serve as surrogates for the unobserved variables of the system. The policy optimization problem is essentially dynamic in nature. It would be unrealistic to assume that policy-makers would make their decisions without taking into consideration the past performance of the economy being controlled. However, the optimization of the process will usually entail the consequence that future policy affects present policy. Note here that there is an intimate link between the dynamic and the stochastic specification.

iii) Econometric models have naturally the tendency to be discrete-time models. In practical applications, the data available are in the form of sampled time series and therefore, the values of the variables are known only at discrete points of time. In a real economy, the agents generally make decisions at each period and the periods are defined by discrete numbers. Decisions occur in sequences. In other words, the process of decision is discrete. This is because the dynamic of the economy is observed only at discrete intervals of time. It appears realistic to assume that the agents act sequentially (at each period), taking into account all the available information. The agents do not modify their decisions continuously. They have only a finite number of possible actions. Empirical models are often built in discrete time setting. Although the models may be perceived in continuous time, their implementation for simulation purposes is most likely to be in discrete-time environment. Numerical solutions necessarily require the reformulation of the problem into a discrete-time finite horizon approximation. In order to estimate the parameters of a continuous time model, it is necessary to relate them to the discretely observed data. The continuous approach is convenient for theorizing but it is less realistic. In the econometric practice, few explanatory variables are continuous. Many of them are dummies, qualitative variables, or counts. Others, even if continuous in nature, are recorded at intervals and can be treated as discrete.

Complex economic systems are generally characterized by vectors of endogenous and exogenous variables. Let $x_t \in \mathbf{R}^q$ be the value of the selected control-related (external) variable at time t (regarded as a strategic instrument-variable of the decision-maker), $y_t \in \mathbf{R}^p$ the system (observable) target internal-variable in t (modelled as a partly or indirectly controlled variable) and $z_t \in \mathbf{R}^r$ an exogenous variable not subjected to the decision-maker' control which is observed outside the system under consideration and so unaffected by the control process. It may be forecasted by the decision-maker but cannot be influenced by him.

Note that x_t is not strictly exogenous, in general the actions being dependent variables on the history and current state of the system. Generally, for employing the input x_t, the decision-maker will incur a certain cost. He will also incur adjustment costs for necessary changes in the inputs. Inevitably, there is an arbitrary element in the choice of control variables and an insufficient variability in the instruments.

We also point out that the target variables are usually a small subset of the total number of endogenous variables in a model.

Whether or not the variable z_t is exogenous depends upon whether or not that variable can be taken as "given" without losing information for the purpose at hand. Specifically, the exogeneity of the variable z_t depends on the parameters of interest of the decision-maker and on the purpose of the model (statistical inference, forecasting, or policy analysis).

The decision-maker chooses an action at any date after he observes the value of the exogenous process $\{z_t\}$ for that date. Thus, the optimal action x_t will depend on the observed value of z_t. The variations in the process z_t over time will therefore result in variations in the process x_t. Complete learning of the true parameter vector will then depend upon whether in the limit the process x_t varies linearly or nonlinearly with the process $\{z_t\}$. Changes in uncertainty about exogenous variables z_t do not lead to changes in the decision-maker's bias. This is not the case when there is a change in uncertainty about policy response parameters.

In general, the size of a model is due more to exogenous variables than to endogenous ones. Due to the stochastic nature of the problem, in practice, it is generally employed a small model.

In what follows, we make the following basic assumptions:

Assumption 1. The evolution of the environment is modeled by the multivariate linear stochastic process:

$$y_t = A_t y_{t-1} + C_t x_t + B_t z_t + D_t + u_t, \qquad t = 1, ..., T.$$

• $\beta_t = (A_t,\ C_t,\ B_t,\ D_t)$ is the endogenous vector of interest parameters, which will be estimated taking into account the information available at time t. It specifies the structure of the model and is utilized to compute the control instrument x_t.

The parameters in the econometric relationships are supposed to vary systematically, according to the information accumulated in the system over time. They are temporarily unstable. The decision-maker knows that shocks will occur in the future, which will need to be counteracted. Thus, future stabilization of the system will be much more effective with more precise estimates of the unknown parameters. The most efficient estimators, in this case, are obtained by using the full information estimation, that is, the whole system is specified and estimated simultaneously.

For optimality reasons, the agent will reestimate the parameters of the model at each time t by taking the feedback effects of learning into account. The uncertainty on the system parameters is thus renewed at each period. The parameter estimates are only revised in response to forecast errors. This regular reevaluation of the parameters certifies that the evolution of the estimated model follows that one of the true process. The process of continuous learning implies an iterative adjustment process and ensures a consistent estimation of the parameters of interest because of the increasingly finer information.

Remark 1. In practice, the decision-maker is confronted with an inherent instability of the system over time. It may be characterized by periods of relatively stable behavior. It exists an intimate link between the dynamic stability of the process and the amount of information it reveals. This is a general phenomenon. Even if we suppose a temporal invariance for the parameter of interest, it cannot ensure the stationarity of the process $\{y_t\}$. Moreover, even if we suppose that $z_t \equiv z$, it is possible that the process be explosive.

Remark 2. In practice, even if the history of the process is longer, the memory of its states is shorter. The predictable impact of y_{t-1} on y_t depends on the degree of persistence parameter A_t. This is the lagged dependent variable which determines the dynamic of the system. It represents the internal force of the system. The smaller the multiplicative slope parameter B_t, the greater needs to have a compensating moving in the control variable x_t. If the parameter on the control variable is large, a small change in the control can cause a much larger change in future state. Time-variation in the unknown parameter B_t implies that the uncertainty regarding this parameter is renewed again and again. A well-known feature of such control problems is the possibility of a trade-off between current control and estimation.

• $u_t \sim i\mathcal{N}(0, \mathbf{\Psi})$, $u_t \in \mathbf{R}^p$, is an exogenous unobserved random shock modelled, for simplicity, by a multivariate normal distribution with zero mean and finite variance $\mathbf{\Psi}$. It symbolizes the unforeseeable elements of the real world, as such the shocks which perturb hazardously the environment, or the errors of measurement. However, nothing forces the data generating process to be stable. The stable distribution is a statistical phenomenon.

Nothing a priori is known about the form of the error distribution. Usually, the form of heteroskedasticity is not specified. The fixed and unknown variance-covariance matrix (as it will typically be in practice) is a non-negative symmetric matrix (but not necessarily of full rank) supposed to depend on an unknown nuisance parameter vector which can be either restricted or not. Unless the model implies a certain form for the covariance matrix, it is desirable to use an estimator that is consistent under the weakest possible conditions (see, Newey & West, 1987, and Andrews, 1991).

There are many statistical ways to generate heteroscedasticity, such as: incorrect data transformation (generally, applied to the dependent variable) or choice of the functional form, changes in the distributions of shocks, the economic behavior and so on. Generally, the heteroscedasticity cannot be completely eliminated. The misspecification of the error covariance matrix is a source of inefficiency in the estimation of the regression parameters (see, ENGLE, 1974).

It is also important to note that the assumption of independence of errors often does not hold or may be of interest to test it. Given that the misspecification is often traduced in practice by autocorrelated errors in the model, it is important to test for autocorrelation before starting the problem of optimization and control.

Assumption 2. The first step in the decision making process is the selection of a feasible objective. Let $\eta \overset{not.}{=} \{y_1^g, y_2^g, ..., y_T^g\}$ be the optimal path desired by the decision-maker. This characterizes his individual preferences on the environment. The targets represents, to a certain extent, the decision-maker's anticipation on the future dynamic of the system, given its backward evolution. They are not uniquely defined. The targets and the dynamics are modelled simultaneously.

Taking into account foreseeable movements in y, the decision-maker will fix some small values for the targets:

$$0 < y_t^g \leq \bar{l}_t' < 1, \qquad t = 1, ..., T$$

where $\bar{l}_t'$ are low optimal bounds. However, it must said that no unique criterion unambiguously determines the values of the targets y_t^g. Assigning extreme values to the targets in order to be sure that the solution of the model always keeps the values of the objectives on one side of the targets, would not be a realistic economic strategy for the decision-maker.

For stochastic control systems, there are many path which the system states may follow given the control and initial data. In this case, the best system performance depends on the information available to the controller at each time t. In the stochastic context, it will exist permanent and significant errors on the control. The negative effect of the system stochasticity will be the control deviation. Since a real-time control process is necessarily discrete, we cannot hope to converge precisely to any target value, but only to some neighborhood of it. In other words, after the process of control is ended, the decision-maker will obtain a stochastic neighbouring-trajectory which is expected to be close to the reference-optimal trajectory. The goal of the control is to maintain the process most of its time near the equilibrium state η. An other explanation why the targets are generally unattainable may be their incompatibility with the state of the system. To a certain extent, the fixed targets are subjective.

It is useful to note that an a priori analysis of the deterministic control problem is often crucial (see, SARGENT, 1987).

Remark 3. When there is no cost on the control, then the decision-maker does not have as objective to follow a fixed optimal trajectory. This is the case of a myopic (or pseudo-optimal) decision behavior.

Assumption 3. The timing of the control is as follows: At each stage t, the agent implements an action x_t which is a stimulus for the system. This is purported to contribute towards equilibrium and stability of the environment. A shock u_t is realized and the agent observes the output y_t (the impulse response) from which he extracts a dynamic signal about the environment trend. The question is: how this signal will influence the behavior of the decision-maker?

This output and the corresponding action provide information on the data generating process. The decision-maker will utilize this output-signal for non strategic learning (he learns the "true" parameters of interest), or strategic learning (specific to a closed-loop strategy) in order to reduce the uncertainty on the future behavior of the system. Note that the uncertainty is reduced only ex-post. The shock u_t will have a persistent effect on y_t, which will disappear gradually over time.

Remark 4. At the end of time period $t-1$, the control rule x_t is applied and the target variable y_t is determined. Thus, the values of x_t and y_t are determined at time $t-1$. During the period $t-1$ to t the environment reacts to these values, as they become generally known, and by the time t arrives, y_t is determined from the state equation. Consequently, there is an apparent instantaneous relationship between x_t and y_t in the state equation.

The time lag between $t-1$ and t is a decision lag, and it should be strongly emphasized that this lag need not correspond to the interval between observations of the environment represented by the data available for analysis. So, the decision lag (or decision time) and the observation interval (or observation time-periods) need not coincide, such that one decision period may equal N observation periods (N could be greater than or less than unity).

Assumption 4. The optimality of the instrument is considered with respect to a global criterion which measures the system deviations $\triangle\, y_t \overset{not.}{=} y_t - y_t^g$, $t = 1, ..., T$. The optimal policy will ensure the minimal deviation between the state of the system and the target (an a priori value or level of aspiration).

Let $W_{[1,T]}(y_1, y_2, ..., y_T)$ be this criterion, supposed twice continuously differentiable and strictly convex at least in the feasible area of the model.

A quadratic objective function may be condisered as a good local approximation of the true preferences, exactly as a model approaches the behavior of the environment around the observed variables. This is a reasonable one since it induces a high penalty for large deviations of the state variable from the target but a relatively small penalty for small deviations. These endogenous deviations exist because of the phenomenon of "learning bunching" (that is, small learning biases are present during some periods while large biases occur during others). Even in cases where the quadratic objective function is not entirely justified, it is still used since it leads to an elegant analytical solution in the linear model case and a computationally feasible numerical solution for the non-linear model. Although the use of an explicit overall criterion is not generally possible, most of the theoretical and practical works accept the hypothesis of a quadratic loss function.

In addition, nothing impedes to suppose that the loss function is additive period by period, on one hand in order to simplify the deduction of the formulas for the optimal instrument and on the other hand, because it will make possible to apply the BELLMAN (1961) optimality principle backwards through time (the decisions being independently stochastic). We therefore consider, in what follows, a global quadratic additively recursive criterion:

$$W_{[1,T]}(y_1, ..., y_T) \overset{def}{=} \sum_{t=1}^{T} W_t(y_t)$$

where W_t is a quadratic asymmetrical loss function, strictly convex and twice differentiable:

$$W_t(y_t) \overset{def.}{=} (y_t - y_t^g)' K_t (y_t - y_t^g) + 2(y_t - y_t^g)' d_t$$

The asymmetry of the criterion in the target values derives from the difference in penalty costs that the decision-maker may attach to errors, depending on whether they are errors of shortfall or errors of overshooting about the target. He is not indifferent with regard to the sign of the system deviations over time. There is an asymmetrical treatment of errors to either side of the target, that is, a positive deviation from a target is not penalized as a negative deviation of the same magnitude.

The criterion for making decisions is a function that puts weight (or measure) on the possible outcomes indicating their desirability or undesirability. The parameters K_t and d_t

allow to weight differently the various loss components. In other words, we will not have equivalent deviations of the target variables during the optimization process. The weights used are, of course, anything but objective, since the deviation of all target variables may be not of the same importance.

In general, the decision for choosing certain parameters K_t and d_t reflects the decision-maker's priorities and also depends on the available amount of information concerning the future development of the system parameters. However, it is unlikely that the decision-maker will be able to assign values to the weights which correctly represent his preferences. The idea is to choose the parameters which yield a smoother control (that is, less fluctuating) and hence a more stable closed-loop system. If the future evolution of the system is unpredictable, then the best weighting matrix K_t which can be chosen is the identity matrix, while the best value for the vector d_t is the unity vector.

At each stage t, the parameters K_t and d_t are updated and new optimal values are chosen in order to satisfy the requirements of the decision-maker. These requirements are based on policy values at each stage, and do not require any direct information about the actual weighting the policymaker may have in his mind.

It is important to note that although widely adopted in the literature, the assumption of the availability of an explicit expression for $W_{[1,T]}$ is not free from critiques, especially when real world problems are considered (see, BOCK & PAULY, 1978). It exists situations where the decision-maker is not able, or is not willing to formulate an explicit criterion function. How sensitive is this assumption and how sensitive to this assumption are the control results rest still a sensible subject in this area.

Assumption 5. At each period t, the agent computes his optimal policy $\widehat{x}_t$ before to know the initial state of the process y_0. He therefore obtains a random optimal policy, conditional to y_0:

$$\widehat{x}_t = \arg\max_{x_t} E_{t-1}[U_t(W_{[1,t]}, \varphi_t) \mid y_0]$$

where $E_{t-1}(\cdot) \overset{not.}{=} E(\cdot \mid I_{t-1})$ represents the operator of conditional expectation based on information available in $t-1$, φ_t is the absolute risk-aversion index at time t based on a truncated history of the process and rational anticipations of the system behavior in the future, and U_t is the agent's anticipative (local) utility function defined by:

$$U_t(W_{[1,t]}, \varphi_t) \overset{def.}{=} \frac{2}{\varphi_t}[\exp(-\frac{\varphi_t}{2} \cdot W_{[1,t]}) - 1]$$

with

$$W_{[1,t]} \overset{def.}{=} \sum_{s=1}^{t} W_s(y_s) \text{ (evolutive loss)}$$

It follows that:

$$-\frac{U''(W_{[1,t]}, \varphi_t)}{U'(W_{[1,t]}, \varphi_t)} = \frac{\varphi_t}{2}$$

where a prime denotes the partial derivative with respect to $W_{[1,t]}$.

Therefore, $\frac{\varphi_t(W_{[1,t]})}{2}$ measures locally (at the point $W_{[1,t]}$) the agent's risk-aversion. Note here that the non-linearity of the utility function is more commonly represented as risk-aversion. For further details, see PROTOPOPESCU, D., 2007.

This is JACOBSON (1973, 1977) the first who has posed in evidence an exponential utility for the problems of stochastic optimal control with (symmetrical) quadratic criterion. Generally speaking, the utility depends on the purposes for which it is developed. It does not exist but for the agent, and thus, it has a subjective character. This is derived from individual preferences. It is very rare in econometrics to be able to fully specify the utility function. No decision-maker has sufficient a priori knowledge to fully specify his utility function. The stochastic disturbance in the system will produce random shocks in the decision-maker's preferences over time. The maximum expected utility solution does not necessarily correspond to a stochastic optimal policy with minimum variance.

In general, the initial state y_0 (a past observation of the dynamic process) is either fixed or randomized (in this latter case, it is possible to have an a priori distribution on y_0 based on the information acquired until the date $t = 0$). Because a real system is always subject to permanent shocks, it is not possible to control its initial state exactly. This will amplify the uncertainty on the system behavior. It will be crucial to achieve a correct treatment of the starting value y_0 and to measure its impact. Small differences in initial conditions can have large effects on long-run outcomes.

Remark 5. A necessary (but not sufficient) condition for the unicity of the instrument is that the number of target variables be inferior to the number of instruments ($p \leq q$).

Remark 6. The only constraints in the above optimization problem are the model equations for each time point in the policy period. However, if nonconvexities arise in the criterion (objective function), this may greatly complicate the search for the optimal control instruments (see, AMMAN and KENDRICK, 1995), and additional constraints for smoothing and bounding the controls may also be present.

Remark 7. It is far from probable that the decision-maker exactly maximizes his utility at each stage of the control. We rather face a nearly optimization behavior, where the control variable is continuously and optimally adjusted to maximize some objective function (see, VAN DE STADT, H. ET AL., 1985, VARIAN, 1990 and LELAND, 1990, among others). The stochastic disturbance in the system will produce random shocks in the decision-maker's preferences over time.

4. Probabilistic Hypotheses on the Acquisition of Information

Given that some random strategies are employed, the stochastic environment must be described by a complete finite probability space $(\Omega, \mathcal{F}, P_\Omega, \mathcal{H})$ endowed with a filtration $\mathcal{H}$ (i.e., an increasing sequence of σ sub-algebras of $\mathcal{F}$) satisfying the "usual" technical conditions.

Denote by $\mathcal{F}$ the σ-algebra of $\mathcal{P}(\Omega)$. P_Ω is the decision-maker's subjective probability measure on Ω ($\mathcal{P}(\Omega) = 1$) and represents the stochastic law of the environment (the agent may be uncertain about the state of the world). In statistical applications, P_Ω is an element of a family of sampling probabilities.

Let $I = \underset{t \leq T}{\cup} I_t$ be the space of all possible "elementary events" in the given environment. It plays the role of Ω. Suppose that the family of events Ω is atomless, that is, that any event but $\emptyset$ (the impossible event) is the union of two exclusive events which are also different from $\emptyset$. This assumption expresses the idea that a refinement of the description of an uncertain environment can always be made. Some additional specific assumptions should be made.

Assumption 6. (**Non-anticipation**): The history of the process, the past actions and the history of the exogenous variables constitute the maximum that can be fully observed and known at a given moment t.

Assumption 7. (**Non-causality**): The future actions cannot affect the current dynamic of the process. The principle of causality requires that the dynamics of the process being such that present or past actions can affect only future outcomes and not vice versa.

Assumption 8. (**Retention of information**): At the date t, the information I_t is I_{t+1}-measurable. Once the information is obtained, this is definitively acquired. In particular, the past actions are memorized. Uncertainty will be solved over time according to a discrete-filtration $\mathcal{H} \overset{not.}{=} \{\mathcal{F}_t \mid t = 0, ..., T\}$ with $\mathcal{F}_T \overset{def.}{=} \mathcal{P}(\Omega)$ and $\mathcal{F}_0 \overset{def.}{=} \{\emptyset, \Omega\}$ almost trivial (meaning that Ω is the only event of non-zero probability in $\mathcal{F}_0$), filtered to the right which respect to the operator of inclusion (i.e., $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$ for all t, and so $\mathcal{F}_t \subset \mathcal{F}_s$ whenever $s \geq t$). In other words, nothing is forgotten, the memory of the process increasing over time.

5. Linear Feedback Optimal Strategy: The Classic Context

The objective of this section is to correct the theoretical results of Van der Ploeg (1984a, 1984b) for the estimation of the feedback optimal strategy in the context of a linear dynamic stochastic environment. We consider here the case where the decision maker's risk-aversion is constant and exogenous by hypothesis. Let φ be the absolute risk-aversion index fixed during the entire control period $[1, T]$.

Proposition 1. (i) Under the hypotheses stated in **Section 2** and **Section 3**, the optimal feedback control equation for the period t is given by:

$$\widehat{x}_t(I_{t-1}, z_t, \beta_t, K_t, d_t, y_t^g) \mid y_0 = G_t \cdot y_{t-1} + g_t, \qquad t = 1, ..., T$$

where:

$$G_t \overset{def.}{=} -(C_t' \widetilde{H}_t C_t)^{-1}(C_t' \widetilde{H}_t A_t)$$

$$g_t \overset{def.}{=} -(C_t' \widetilde{H}_t C_t)^{-1} C_t' [\widetilde{H}_t (B_t z_t + D_t) - (I_p - \varphi K_t (\mathbf{\Psi}^{-1} + \varphi \cdot H_t)^{-1}) h_t]$$

$$\widetilde{H}_t \overset{not.}{=} K_t - \varphi \cdot H_t M_t^{-1}(\varphi) H_t, \quad M_T(\varphi) \overset{not.}{=} \mathbf{\Psi}^{-1} + \varphi \cdot H_T$$

(ii) It exists the following backward recurrences ($t = T, T-1, ..., 1$):

$$H_{t-1} = K_{t-1} + (A_t + C_t G_t)' \widetilde{H}_t (A_t + C_t G_t)$$

$$h_{t-1} = K_{t-1} y_{t-1}^g - (A_t + C_t G_t)' [\widetilde{H}_t (C_t g_t + B_t z_t + D_t) - (I_p - \varphi \cdot K_t (\mathbf{\Psi}^{-1} + \varphi \cdot H_t)^{-1}) h_t]$$

with initial conditions

$$H_T \overset{not.}{=} K_T \text{ and } h_T \overset{not.}{=} K_T y_T^g - d_T$$

Proof. The dynamic programming problem is approached in finite discrete-time and uncertain future. Given the assumptions of non-anticipation, retention of information, and additivity for the global loss function $W_{[1,T]}$, the multiperiod stochastic optimization problem (T subperiods) can be decomposed into a sequence of local optimization problems (see, Bellman, R., 1961):

$$\underset{x_1,...,x_T}{\arg\max} E_0 U_T(W_{[1,T]}, \varphi) = \underset{x_1(\cdot)}{\arg\max} E_0 [\underset{x_2(\cdot)}{\arg\max} E_1 (... \underset{x_T(\cdot)}{\arg\max} E_{T-1} U_T(W_{[1,T]}, \varphi)]$$

where $U_T(W_{[1,T]}, \varphi) \overset{def}{=} \frac{2}{\varphi}\{\exp[-(\frac{\varphi}{2}) W_{[1,T]}] - 1\}$ represents the agent's utility function at time T, $E_{t-1}(\cdot) \overset{not.}{=} E(\cdot \mid I_{t-1})$ is the operator of conditional expectation based on the information available to the agent at time $t-1$, and φ defines the absolute risk-aversion index fixed for the entire period of control.

It comes to maximize period by period (sequential decision problem), working every time conditionally to the information acquired. The optimal policy is computed step by step starting from x_T towards x_1 (backward through time). We first consider the decision problem for the last period T, given all the information available at the end of period $T-1$. We can write:

$$E_{T-1}U_T(W_{[1,T]}(y_T),\varphi) = \frac{2}{\varphi}E_{T-1}[\exp\{-(\frac{\varphi}{2})W_T(y_T)\}\exp\{-(\frac{\varphi}{2})\sum_{t=1}^{T-1}W_t(y_t)\}-1]$$

Because the last exponentiel does not depend on x_T, we have:

$$\widehat{x}_T = \underset{x_T}{\arg\max} E_{T-1}U_T(W_{[1,T]}(y_T),\varphi) = \underset{x_T}{\arg\max} E_{T-1}[\exp\{-(\frac{\varphi}{2})W_T(y_T)\}]$$

The assumption of rational expectations makes the problem difficult because the expected value of a non-linear function is not generally the non-linear function of the expected value of the random variable. Under appropriate regularity conditions, we can interchange the order of integration and differentiation, that is, we can differentiate within the conditional expectation operator.

For the computation of $E_{T-1}[\exp\{-(\frac{\varphi}{2})W_T(y_T)\}]$ (which is supposed to exist), we proceed as follows:

$$E_{T-1}[\exp\{-(\frac{\varphi}{2})W_T(y_T)\}] = E_{T-1}[\exp\{-(\frac{\varphi}{2})\cdot(\triangle\, y_T'K_T\triangle\, y_T + 2\triangle\, y_T'd_T)\}]$$

$$= E_{T-1}[\exp\{-(\frac{\varphi}{2})(y_T'H_Ty_T - 2y_T'h_T + f_T)\}]$$

where:

$$\triangle\, y_T \overset{not.}{=} y_T - y_T^g,\quad H_T \overset{def.}{=} K_T,\quad h_T \overset{not.}{=} K_Ty_T^g - d_T,\quad f_T \overset{not.}{=} y_T^{g\prime}(h_T - d_T)$$

Substituting $A_Ty_{T-1} + C_Tx_T + B_Tz_T + D_T + u_T$ for y_T, we obtain:

$$V_T \overset{not.}{=} E_{T-1}[\exp\{-(\frac{\varphi}{2})W_T(y_T)\}] = E_{T-1}[\exp\{-(\frac{\varphi}{2})\cdot y_T'H_Ty_T + \varphi\cdot y_T'h_T - (\frac{\varphi}{2})f_T\}]$$

$$= E_{T-1}[\exp\{-(\frac{\varphi}{2})\cdot u_T'H_Tu_T - \varphi\cdot u_T'[K_T\cdot(A_Ty_{T-1}+C_Tx_T+B_Tz_T+D_T) - h_T] - (\frac{\varphi}{2})f_T\}\cdot$$

$$\exp\{-(\frac{\varphi}{2})\cdot\{(A_Ty_{T-1}+C_Tx_T+B_Tz_T+D_T)'[K_T(A_Ty_{T-1}+C_Tx_T+B_Tz_T+D_T) - 2h_T]\}\}]$$

$$= E_{T-1}[\exp(\omega_2(u_T))]\exp\omega_1(I_{T-1},x_T,z_T,\beta_T,K_T,d_T,y_T^g)$$

$$= \exp\omega_1(I_{T-1},x_T,z_T,\beta_T,K_T,d_T,y_T^g)\int_{R^p}(2\pi)^{-\frac{p}{2}}\mid\det\mathbf{\Psi}\mid^{-\frac{1}{2}}\exp\{-\frac{1}{2}\widetilde{u}_T'\mathbf{\Psi}^{-1}\widetilde{u}_T\}\exp\omega_2(\widetilde{u}_T)\,d\widetilde{u}_T$$

with $\omega_2(\widetilde{u}_T)$ a quadratic function in $\widetilde{u}_T$.

One can write:

$$I \overset{not.}{=} \int_{R^p}(2\pi)^{-\frac{p}{2}}\mid\det\mathbf{\Psi}\mid^{-\frac{1}{2}}\exp\{-\frac{1}{2}\widetilde{u}_T'\mathbf{\Psi}^{-1}\widetilde{u}_T\}\exp\omega_2(\widetilde{u}_T)\,d\widetilde{u}_T$$

$$= \int_{R^p}(2\pi)^{-\frac{p}{2}}\mid\det\mathbf{\Psi}\mid^{-\frac{1}{2}}\exp\{-\frac{1}{2}\widetilde{u}_T'(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)\widetilde{u}_T + (\text{linear in }\widetilde{u}_T')\}d\widetilde{u}_T$$

$$= \mid\det(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)\mid^{-\frac{1}{2}}\mid\det\mathbf{\Psi}\mid^{-\frac{1}{2}}\int_{R^p}(2\pi)^{-\frac{p}{2}}\mid\det(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)\mid^{\frac{1}{2}}\exp\omega_3(\widetilde{u}_T)\,d\widetilde{u}_T$$

with $\omega_3(\widetilde{u}_T)$ a quadratic function in $\widetilde{u}_T$. Now, we find $\overline{u}_T \in R^p$ such that:

$$\omega_3(\widetilde{u}_T) = -\frac{1}{2}(\widetilde{u}_T - \overline{u}_T)'(\mathbf{\Psi}^{-1} + \varphi \cdot H_T)(\widetilde{u}_T - \overline{u}_T) + \text{independent of } \widetilde{u}_T$$

By consequence, we must impose the following equality:

$$-\frac{1}{2}\widetilde{u}'_T(\mathbf{\Psi}^{-1} + \varphi \cdot H_T)\widetilde{u}_T - \varphi \cdot \widetilde{u}'_T[K_T \cdot (A_T y_{T-1} + C_T x_T + B_T z_T + D_T) - h_T] - (\frac{\varphi}{2})f_T$$

$$= -\frac{1}{2}\widetilde{u}'_T(\mathbf{\Psi}^{-1} + \varphi \cdot H_T)\widetilde{u}_T + \widetilde{u}'_T(\mathbf{\Psi}^{-1} + \varphi \cdot H_T)\overline{u}_T - \frac{1}{2}\overline{u}'_T(\mathbf{\Psi}^{-1} + \varphi \cdot H_T)\overline{u}_T$$
$$+\text{independent of } \widetilde{u}_T$$

It follows that:

$$\text{independent of } \widetilde{u}_T = \frac{1}{2}\overline{u}'_T(\mathbf{\Psi}^{-1} + \varphi \cdot H_T)\overline{u}_T - (\frac{\varphi}{2})f_T \overset{not.}{=} \omega_4(\overline{u}_T)$$

and

$$-\varphi \cdot \widetilde{u}'_T[K_T \cdot (A_T y_{T-1} + C_T x_T + B_T z_T + D_T) - h_T]$$
$$= \widetilde{u}'_T(\mathbf{\Psi}^{-1} + \varphi \cdot H_T)\overline{u}_T$$

that is,

$$\overline{u}_T = -\varphi(\mathbf{\Psi}^{-1} + \varphi \cdot H_T)^{-1}[K_T \cdot (A_T y_{T-1} + C_T x_T + B_T z_T + D_T) - h_T]$$

Thus, the integral becomes:

$$I = \ |\det(\mathbf{\Psi}^{-1} + \varphi \cdot H_T)|^{-\frac{1}{2}} |\det \mathbf{\Psi}|^{-\frac{1}{2}} \exp\{\frac{1}{2}\overline{u}'_T(\mathbf{\Psi}^{-1} + \varphi \cdot H_T)\overline{u}_T - (\frac{\varphi}{2})f_T\} \cdot$$

$$\int_{R^p} (2\pi)^{-\frac{p}{2}} |\det(\mathbf{\Psi}^{-1} + \varphi \cdot H_T)|^{\frac{1}{2}} \exp\{-\frac{1}{2}(\widetilde{u}_T - \overline{u}_T)'(\mathbf{\Psi}^{-1} + \varphi \cdot H_T)(\widetilde{u}_T - \overline{u}_T)\}d\widetilde{u}_T$$

The last integral is equal to 1 because the integrand is the probability density function of a p-dimensioanl normal random variable:

$$\widetilde{u}_T \sim \mathcal{N}(\overline{u}_T, (\mathbf{\Psi}^{-1} + \varphi \cdot H_T)^{-1})$$

One can write:

$$|\det(\mathbf{\Psi}^{-1} + \varphi \cdot H_T)|^{-\frac{1}{2}} |\det \mathbf{\Psi}|^{-\frac{1}{2}}$$
$$= \ |\det[\mathbf{\Psi}^{-1}(I_p + \varphi \cdot \mathbf{\Psi} \cdot H_T)\mathbf{\Psi}]|^{-\frac{1}{2}} \ = \ |\det(I_p + \varphi \cdot \mathbf{\Psi} \cdot H_T)|^{-\frac{1}{2}}$$

If we replace $\overline{u}_T$ by its value, we find without difficulty:

$$I = \ |\det(I_p + \varphi \cdot \mathbf{\Psi} \cdot H_T)|^{-\frac{1}{2}} \exp\{-(\frac{\varphi}{2})[K_T(A_T y_{T-1} + C_T x_T + B_T z_T + D_T) - h_T]'$$

$$\cdot - (\varphi) \cdot (\mathbf{\Psi}^{-1} + \varphi \cdot H_T)^{-1} \cdot [K_T(A_T y_{T-1} + C_T x_T + B_T z_T + D_T) - h_T] - (\frac{\varphi}{2})f_T\}$$

$$= \ |\det(I_p + \varphi \cdot \mathbf{\Psi} \cdot H_T)|^{-\frac{1}{2}} \exp \omega_4(I_{T-1}, x_T, z_T, \beta_T, K_T, d_T, y_T^g)$$

By consequence, we have:

$$V_T \overset{not.}{=} E_{T-1}[\exp\{-(\frac{\varphi}{2})W_T(y_T)\}] = \exp \omega_1(I_{T-1}, x_T, z_T, \beta_T, K_T, d_T, y_T^g) \cdot I$$

$$= \ |\det(I_p+\varphi\cdot\mathbf{\Psi}\cdot H_T)\,|^{-\frac{1}{2}}\exp\{\omega_1(I_{T-1},x_T,z_T,\beta_T,K_T,d_T,y_T^g)+\omega_4(I_{T-1},x_T,z_T,\beta_T,K_T,d_T,y_T^g)\}$$

$$= \ |\det(I_p+\varphi\cdot\mathbf{\Psi}\cdot H_T)\,|^{-\frac{1}{2}}\exp\omega_5(I_{T-1},x_T,z_T,\beta_T,K_T,d_T,y_T^g)$$

After several algebraic manipulations, we find:

$$\omega_5(I_{T-1},x_T,z_T,\beta_T,K_T,d_T,y_T^g)=-(\frac{\varphi}{2})(A_Ty_{T-1}+C_Tx_T+B_Tz_T+D_T)'$$
$$\cdot[K_T-\varphi H_T(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)^{-1}H_T](A_Ty_{T-1}+C_Tx_T+B_Tz_T+D_T)$$
$$+\varphi(A_Ty_{T-1}+C_Tx_T+B_Tz_T+D_T)'[I_p-\varphi\cdot K_T(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)^{-1}]h_T$$
$$+(\frac{\varphi}{2})\varphi h_T'(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)^{-1}h_T-(\frac{\varphi}{2})f_T$$
$$=-(\frac{\varphi}{2})(y_{T-1}'A_T'+x_T'C_T'+z_T'B_T'+D_T')\widetilde{H}_T(A_Ty_{T-1}+C_Tx_T+B_Tz_T+D_T)$$
$$+\varphi(y_{T-1}'A_T'+x_T'C_T'+z_T'B_T'+D_T')[I_p-\varphi\cdot K_T(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)^{-1}]h_T$$
$$+(\frac{\varphi}{2})\varphi h_T'(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)^{-1}h_T-(\frac{\varphi}{2})f_T=-(\frac{\varphi}{2})[y_{T-1}'A_T'\widetilde{H}_TC_Tx_T$$
$$+x_T'C_T'\widetilde{H}_T(A_Ty_{T-1}+B_Tz_T+D_T)+x_T'C_T'\widetilde{H}_TC_Tx_T+(z_T'B_T'+D_T')\widetilde{H}_TC_Tx_T]$$
$$+\varphi x_T'C_T'[I_p-\varphi\cdot K_T(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)^{-1}]h_T+\text{independent of }x_T$$

where:

$$\widetilde{H}_T\overset{not.}{=}K_T-\varphi\cdot H_TM_T^{-1}(\varphi)H_T$$
$$M_T(\varphi)\overset{not.}{=}\mathbf{\Psi}^{-1}+\varphi\cdot H_T=\mathbf{\Psi}^{-1}(\varphi\mathbf{\Psi}+H_T^{-1})H_T$$

Using the well-known formulas for the derivatives of matricial functions, the first order condition in x_T writes:

$$-(\frac{\varphi}{2})C_T'\widetilde{H}_TA_Ty_{T-1}-(\frac{\varphi}{2})C_T'\widetilde{H}_T(A_Ty_{T-1}+B_Tz_T+D_T)-\varphi C_T'\widetilde{H}_TC_Tx_T$$
$$-(\frac{\varphi}{2})C_T'\widetilde{H}_T\,(B_Tz_T+D_T)+\varphi C_T'[I_p-\varphi\cdot K_T(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)^{-1}]h_T=0\overset{(\varphi\neq 0)}{\Leftrightarrow}$$
$$-C_T'\widetilde{H}_TA_Ty_{T-1}-C_T'\widetilde{H}_T\,(B_Tz_T+D_T)+C_T'[I_p-\varphi\cdot K_T(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)^{-1}]h_T=C_T'\widetilde{H}_TC_Tx_T$$

It follows that:

$$\widehat{x}_T(I_{T-1},z_T,\beta_T,K_T,d_T,y_T^g)=G_T\cdot y_{T-1}+g_T \quad (1)$$
$$G_T\overset{not.}{=}-(C_T'\widetilde{H}_TC_T)^{-1}(C_T'\widetilde{H}_TA_T) \quad (2)$$
$$g_T\overset{not.}{=}-(C_T'\widetilde{H}_TC_T)^{-1}C_T'[\widetilde{H}_T\,(B_Tz_T+D_T)-(I_p-\varphi K_T(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)^{-1})h_T] \quad (3)$$

The expected utility level for the period T is obtained by substituting for x_T in $\overline{U}_T$:

$$\widehat{V}_T\ \overset{not.}{=}\ |\det(I_p+\varphi\cdot\mathbf{\Psi}\cdot H_T)\,|^{-\frac{1}{2}}\exp\omega_5(I_{T-1},\widehat{x}_T,z_T,\beta_T,K_T,d_T,y_T^g)$$

$$= \ |\det(I_p+\varphi\cdot\mathbf{\Psi}\cdot H_T)\,|^{-\frac{1}{2}}\exp\{-(\frac{\varphi}{2})[(A_T+C_TG_T)\cdot y_{T-1}+C_Tg_T+B_Tz_T+D_T]'\cdot$$
$$\cdot\widetilde{H}_T[(A_T+C_TG_T)\cdot y_{T-1}+C_Tg_T+B_Tz_T+D_T]$$
$$+\varphi[(A_T+C_TG_T)\cdot y_{T-1}+C_Tg_T+B_Tz_T+D_T]'[I_p-\varphi\cdot K_T(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)^{-1}]h_T$$

$$+(\frac{\varphi}{2})\varphi h_T'(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)^{-1}h_T-(\frac{\varphi}{2})f_T\}$$

$$= \ |\det(I_p+\varphi\cdot\mathbf{\Psi}\cdot H_T)|^{-\frac{1}{2}}\exp\{-(\frac{\varphi}{2})[y_{T-1}'(A_T+C_TG_T)'\widetilde{H}_T(A_T+C_TG_T)\cdot y_{T-1}$$

$$+2y_{T-1}'(A_T+C_TG_T)'\widetilde{H}_T(C_Tg_T+B_Tz_T+D_T)+(C_Tg_T+B_Tz_T+D_T)'\widetilde{H}_T(C_Tg_T+B_Tz_T+D_T)$$

$$-2y_{T-1}'(A_T+C_TG_T)'(I_p-\varphi\cdot K_T(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)^{-1})h_T-2(C_Tg_T+B_Tz_T+D_T)'$$

$$\cdot(I_p-\varphi\cdot K_T(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)^{-1})h_T-\varphi h_T'(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)^{-1}h_T+f_T]\}$$

Now, we include the period $T-1$ in our optimization problem. We have:

$$\widehat{x}_{T-1}\overset{def.}{=}\underset{x_{T-1}}{\arg\max}E_{T-2}\{E_{T-1}U_T(W_{[1,T]}(y_T),\varphi)\}=$$

$$=\ \underset{x_{T-1}}{\arg\max}E_{T-2}[E_{T-1}[\exp\{-(\frac{\varphi}{2})W_T(y_T)\}\exp\{-(\frac{\varphi}{2})\sum_{t=1}^{T-1}W_t(y_t)\}-1]]$$

$$=\ \underset{x_{T-1}}{\arg\max}E_{T-2}[E_{T-1}[\exp\{-(\frac{\varphi}{2})W_T(y_T\,(\widehat{x}_T))\}\cdot\exp\{-(\frac{\varphi}{2})W_{T-1}(y_{T-1})\}]]$$

$$=\ \underset{x_{T-1}}{\arg\max}E_{T-2}[\exp\{-(\frac{\varphi}{2})W_{T-1}(y_{T-1})\}E_{T-1}[\exp\{-(\frac{\varphi}{2})W_T(y_T\,(\widehat{x}_T))\}]]$$

The expected utility level for the two last sub-periods is therefore:

$$V_{T-1}\overset{not.}{=}E_{T-2}[\exp\{-(\frac{\varphi}{2})W_{T-1}(y_{T-1})\}\widehat{V}_T]\ =\ |\det(I_p+\varphi\cdot\mathbf{\Psi}\cdot H_T)|^{-\frac{1}{2}}$$

$$\cdot E_{T-2}[\exp\{-(\frac{\varphi}{2})[y_{T-1}'K_{T-1}y_{T-1}-2y_{T-1}'K_{T-1}y_{T-1}^g+y_{T-1}^{g\prime}K_{T-1}y_{T-1}^g$$

$$+y_{T-1}'(A_T+C_TG_T)'\widetilde{H}_T(A_T+C_TG_T)\cdot y_{T-1}+2y_{T-1}'(A_T+C_TG_T)'\widetilde{H}_T(C_Tg_T+B_Tz_T+D_T)$$

$$-2y_{T-1}'(A_T+C_TG_T)'(I_p-\varphi\cdot K_T(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)^{-1})h_T$$

$$+(C_Tg_T+B_Tz_T+D_T)'\widetilde{H}_T(C_Tg_T+B_Tz_T+D_T)-2(C_Tg_T+B_Tz_T+D_T)'$$

$$\cdot(I_p-\varphi\cdot K_T(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)^{-1})h_T-\varphi h_T'(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)^{-1}h_T+f_T]\}]$$

$$\overset{not.}{=}|\det(I_p+\varphi\cdot\mathbf{\Psi}\cdot H_T)|^{-\frac{1}{2}}\cdot E_{T-2}[\exp\{-(\frac{\varphi}{2})[y_{T-1}'H_{T-1}y_{T-1}-2y_{T-1}'h_{T-1}+f_{T-1}]\}]$$

where, by identification, we obtain the following recurrences:

$$H_{T-1}=K_{T-1}+(A_T+C_TG_T)'\widetilde{H}_T(A_T+C_TG_T)$$

$$h_{T-1}=K_{T-1}y_{T-1}^g-(A_T+C_TG_T)'\,[\widetilde{H}_T(C_Tg_T+B_Tz_T+D_T)$$

$$-(I_p-\varphi\cdot K_T(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)^{-1})h_T]$$

$$f_{T-1}=y_{T-1}^{g\prime}K_{T-1}y_{T-1}^g+(C_Tg_T+B_Tz_T+D_T)'\widetilde{H}_T(C_Tg_T+B_Tz_T+D_T)$$

$$-2(C_Tg_T+B_Tz_T+D_T)'\cdot(I_p-\varphi\cdot K_T(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)^{-1})h_T+f_T-\varphi h_T'(\mathbf{\Psi}^{-1}+\varphi\cdot H_T)^{-1}h_T$$

The solution for $\widehat{x}_{T-1}$ will be identical with (1) with T replaced by $T-1$, where G_{T-1} and g_{T-1} are defined by (2) and (3), respectively, with a similar change in time subscripts. We can thus apply a backward induction in time ($t=T,T-1,...,1$) in order to find the decision-maker's optimal strategy for all sub-periods. At the end of this process, we find $\widehat{x}_1=G_1y_0+g_1$ as the optimal policy for the first period and the associated maximum expected utility for all periods. The determination of the optimal $\widehat{x}_1$ depends on the method of forward-looking which is used in the optimality for the future decisions. We cannot obtain an optimal policy for the first period if we do not know its behavior in the future.

The matrices G_t are obtained by solving the matrix equations:

$$G_t = -(C_t' \widetilde{H}_t C_t)^{-1} (C_t' \widetilde{H}_t A_t)$$

$$H_{t-1} = K_{t-1} + (A_t + C_t G_t)' \widetilde{H}_t (A_t + C_t G_t)$$

backward in time with initial condition $H_T = K_T$. Also, the vectors g_t are obtained by solving:

$$g_t = -(C_t' \widetilde{H}_t C_t)^{-1} C_t' [\widetilde{H}_t (B_t z_t + D_t) - (I_p - \varphi K_t (\mathbf{\Psi}^{-1} + \varphi \cdot H_t)^{-1}) h_t]$$

$$h_{t-1} \overset{not.}{=} K_{t-1} y_{t-1}^g - (A_t + C_t G_t)' [\widetilde{H}_t (C_t g_t + B_t z_t + D_t) - (I_p - \varphi \cdot K_t (\mathbf{\Psi}^{-1} + \varphi \cdot H_t)^{-1}) h_t]$$

backward in time with initial condition $h_T = K_T y_T^g - d_T$.

These formulas correct those obtained by Van der Ploeg (1984a, 1984b) and utilized in the literature related to stochastic feedback optimal control (see, Jacobson, 1973, Karp, 1987 and Whittle, 1981, 1989, 1990, among others). In particular, for $\varphi = 0$, we obtain a correction of the classic results obtained by Chow (1973, 1976a, 1976b, 1977, 1978, 1981, 1993) in the risk-neutral context.

Note that in the stationary case (very often utilized for econometric purposes), when all system parameters as well as the parameters of the optimal feedback control equation are supposed to be time-invariant, the constraints for the matrices (G, g) will be modified with respect to the case $\varphi = 0$. We will have not only a dependence on φ (which allows for estimating or testing its values), but also a more complicated discrete Riccati non-linear stochastic matrix equation (which must be iterated back in time in order to obtain the optimal policy in any period). The resulting decision rules will have recursive and adaptive properties (e.g., Kalman filters) in the sense that they can be sequentially updated by incorporating posterior knowledge about means and variances when the information sequence is increasing.

It is possible that differences between ex-ante decisions and ex-post results (in other words, between ex-ante and ex-post optimality) exist. What was in a decision-maker's ex-ante best interest is not necessarily in his ex-post best one.

Even if the linear approximation is only roughly, we can however implement a feedback strategy for a closed-loop dynamic process which is sufficiently good, on one hand for obtaining the evaluation of the policy for the first period and on the other hand for the actual implementation in the future.

6. Endogenous Risk-Aversion

The uncertainty always attends the risk. In the real world, the decision-maker is confronted with multiple risks (generally, different phenomena are characterized by different risks). His decision is not made independently; it is made together with other decisions that place the decision-maker in risky situations. Decisions made to avoid even partially one source of risk may be affected by the presence of others.

It is well-known that economic agents behave on average risk-neutral for small and repeated decisions, but the most common attitude of economic agents in all important decision making problems is one generated by risk-aversion (they prefer the expected value of the risk to the risk itself). Such a behavior characterizes most decision-makers, at least for large gains or large losses. An agent who expects in the future high deviations from the fixed targets can be considered to be risk-averse.

In general, the aversion is associated with increasing uncertainty while the uncertainty is naturally associated with incomplete information about future behavior of the system. One can interpret the risk as the decision-maker's degree of confidence in the future. It decreases with the uncertainty. Traditionally, the risk-aversion is equivalent to the concavity of the agent's utility function or a decreasing marginal utility. However, this is just a way of expressing risk-averse preferences.

In the literature on risk, one generally assume that uncertainty is uniformly distributed over the entire working horizon, when the absolute risk-aversion index is negative and constant. From this perspective, the risk is totally exogenous, and thus independent of endogenous risks. The classic procedure is "myopic" with regard to potential changes in the future behavior of the agent due to inherent fluctuations of the system over time. The traditional measures of risk-aversion are generally too weak for making comparisons between risky situations. This can be highlighted in concrete problems in finance and insurance, context for which the ARROW-PRATT measures (in the small) give ambiguous results (see, ROSS, 1981). We extend the ARROW-PRATT approach (1964, 1971A, 1971B), which takes into account only attitudes towards small exogenous risks, by integrating in the analysis potentially high endogenous risks which are at least partially controllable by the agent. This point of view has strong implications on the agent's adaptive behavior towards risk in an evolving environment.

In any uncertain environment, the decision-maker must form expectations. When his uncertainty is high it may well be that there is a discrepancy between what he expects and reality. In an noisy environment, the expectations may be disappointed.

The decision-maker can influence the likelihood of the environment states by using a reinforcement learning strategy. We say that he is not myopic in the sense of expecting. A myopic behavior leads to an important bias in the controls and targets variables. Future anticipations play an important role in how the agent will decide what strategic actions and optimal risk to take. Agent's behavior will depend on forecasts of future environmental state. If the data generating process changes in ways not anticipated by the model, then forecasts lose accuracy. Without uncertainty, the distinction between the present and the future is confused and there is no anticipation.

Suppose that the agent is a strategic decision-maker. He thinks about the future. Depending on the way the agent perceives future outcomes, both risk sensitivity and optimal decisions will be affected during the process of optimization and control. The forecast is updated each time as new observation becomes available. The agent's rationality is characterized by the fact that the sequence of updated forecasts will converge to the equilibrium of the system.

Different forecasts are obtained from different information structure. There are several sources of forecast uncertainty, including parameter non-constancy, estimation uncertainty, variable uncertainty, innovation uncertainty and model misspecification.

A correct evaluation of the past is crucial for making optimal predictions in the future. This is necessary for an optimal assessment of the agent's risk-aversion over time. Fluctuations in the system target variable generate a time-varying risk-aversion for the decision-maker during the period of control.

We make the following useful notations:

$$S_{t,\ p\_d} \overset{not.}{=} \parallel y_{t-1} - y^g_{t-1} \parallel^2 + ... + \parallel y_{t-k_1} - y^g_{t-k_1} \parallel^2$$
(the sum of squared past deviations at time t)

$$S_{t,\ a\_f\_d} \overset{not.}{=} \parallel y^a_{t|I_t} - y^g_t \parallel^2 + ... + \parallel y^a_{t+k_2|I_{t+k_2}} - y^g_{t+k_2} \parallel^2$$
(the sum of squared anticipated future deviations at time t)

$$S_{t,\ w\_p\_d} \overset{not.}{=} \parallel y_{t-1} - y^g_{t-1} \parallel^2 L_{t-1} + ... + \parallel y_{t-k_1} - y^g_{t-k_1} \parallel^2 L_{t-k_1}$$
(the weighted sum of squared past deviations at time t)

$$S_{t,\ w\_a\_f\_d} \overset{not.}{=} \parallel y^a_{t|I_t} - y^g_t \parallel^2 \overline{L}_t + ... + \parallel y^a_{t+k_2|I_{t+k_2}} - y^g_{t+k_2} \parallel^2 \overline{L}_{t+k_2}$$

(the weighted sum of squared anticipated future deviations at time t)

where y^g_{t+i} $(i = 0, ..., k_2)$ represent fixed targets in the future (taking into account foreseeable movements in y), $y^a_{t+i|I_{t+i}}$ $(i = 0, ..., k_2)$ are expected values of the target variable at time $t + i$ based on non-decreasing endogenous information sets I_{t+i} and L_{t-j_1} $(j_1 = 1, ..., k_1)$, $\overline{L}_{t+j_2}$ $(j_2 = 0, ..., k_2)$ are weighting scalars attached to the system deviations (in the past and future) with respect to the equilibrium path η.

Econometric forecasting is a very useful instrument of the decision-maker. In a real decision making problem, the forecast must be as accurate and efficient as possible. A necessary preliminary step for the decision-maker in order to optimally choose the target path, is to make some a priori expectations on the future evolution of the system based on its past performances. A question arises. Are the current and past values of the process y_t sufficient to forecast y_{t+k} $(k = 1, ..., k_2)$?

These ex-ante expectations refer to those which held prior to the acquisition of information and generally imply a discrete-time process of tatonnement. They must be unique and in accord with the agent's observations and generally are dependent on the initial value of the state variable. The more they are distant in time, the more they are difficult to assess. (due to the extreme uncertainty of the far future). The ex-ante and ex-post forecast errors will be viewed, in this context, as some indicators of uncertainty and of difficulty of the decision making.

We are now in a position to give a definition of the agent's risk-aversion index by taking into account past performances of the system (a truncated history) and rational anticipations of the system behavior in the future.

Definition. Using t to denote time, the absolute risk-aversion index φ_t evolves according to:

$$\varphi_t \overset{def.}{=} \frac{S_{t,\ w\_p\_d} + S_{t,\ w\_a\_f\_d}}{\sqrt{(S_{t,\ p\_d} + S_{t,\ a\_f\_d})^2 + l}},\ t = 1, ..., T$$

where $l \geq 1$ is a positive integer which characterizes the agent's type, and:

$$1 \leq k_1 < T,\ k_2 \geq 0,\ 1 \leq k_1 + k_2 \leq T - 1$$
$$-1 < L_{t-1} \leq ... \leq L_{t-k_1} \leq 0,\ \ -1 < \overline{L}_t \leq ... \leq \overline{L}_{t+k_2} \leq 0$$

The weights may differ across individuals. They are updated each time as new observation becomes available. The decision-maker gives a higher importance to the past and future deviations which are closer to the moment of implementation of a new optimal action. Smaller the weight is, higher is the importance given by the agent to the system deviation from his local objective.

Given the potential destabilizing role of a long memory, the agent will include in the analysis only a limited history of the process. Distant past observations might increase significantly the biais of the estimators in the econometric model. Generally, they provide an imprecise signal for the decision-maker.

In general, it exists an arbitrary element as regards the choice of the backward lag k_1. The objective is to find the better compromise between fit and complexity. The larger the forward lag k_2 is, the more the prediction error increases. Distant forecasts are difficult to formulate due to unpredictable external disturbances which generally affect the system performance.

It is only by taking into account both, the past and the expected future, that the agent can optimally evaluate the risk in an evolving environment. It allows for a better risk allocation at each period of control. A mixing of objectivity and subjectivity will always characterize the agent's degree of risk-aversion. The complexity of this mixing is given by the changing environment design and the agent's typology.

The higher the degree of risk-aversion at time t, the lower the absolute risk-aversion index φ_t. It may be possible that $\varphi_t \approx 0$ for $t = 1, ...T$. In this case, the agent is characterized by an almost null risk-aversion during the period of control. An other possibility is to obtain $\varphi_{t_1} = \varphi_{t_2}$ for $t_1 \neq t_2$, that is, a constant risk-aversion for distinct periods of time. When $\varphi_t \approx \varphi$ (costant) for $t = 1, ..., T$, then the agent will have an almost constant risk-aversion during the entire planning horizon $[1, T]$. It is important to distinguish between local risk-aversion (at each period t) and global risk-aversion over the entire period of control.

The experimental evidence shows that individuals overweight extreme events. Let $\varphi_{\min}$ be an optimal risk-aversion threshold fixed by the decision-maker before starting the control and for the entire working horizon $[1, T]$. The objective is not to exceed this fixed threshold, if not the agent becomes excessively risk-averse for the current period of control, being characterized by an extreme pessimism. This optimal threshold is chosen such that it offers the best characterization of the agent's type. An agent with a higher (smaller) risk-aversion before starting the control will choose a smaller (higher) threshold $\varphi_{\min}$. If φ_t characterizes the local risk-aversion of the agent (at time t), $\varphi_{\min}$ will characterize his global risk-aversion (over the whole). For further details, see Protopopescu, D., 2007.

7. Linear Optimal Feedback Strategy Sensitive to Controlled Endogenous Risk-Aversion

In this section, we improve the formulas obtained for the optimal feedback control rules in the context of a constant exogenous risk-avesrion, by considering here the more realistic case of a time-varying endogenous risk-aversion subjected to the control of the decision-maker.

Proposition 2. Under the hypotheses stated in **Section 2** and **Section 3**, the linear feedback control equations for a rational decision-maker characterized by endogenous risk-aversion are given by:

$$\widehat{x}_t(I_{t-1}, z_t, \beta_t, K_t, d_t, y_t^g) \mid y_0 = \overline{G}_t \cdot y_{t-1} + \overline{g}_t, \qquad t = 1, ..., T$$

with the following optimal reaction coefficients:

$$\overline{G}_t \overset{not.}{=} -(C_t'\overline{H}_t C_t)^{-1}(C_t'\overline{H}_t A_t)$$

$$\overline{g}_t \overset{not.}{=} -(C_t'\overline{H}_t C_t)^{-1} C_t'[\overline{H}_t (B_t z_t + D_t) - (I_p - \varphi_t K_t(\mathbf{\Psi}^{-1} + \varphi_t \cdot H_t)^{-1}) h_t]$$

$$\overline{H}_t \overset{not.}{=} K_t - \varphi_t \cdot H_t M_t^{-1}(\varphi_t) H_t, \qquad M_t(\varphi_t) \overset{not.}{=} \mathbf{\Psi}^{-1} + \varphi_t \cdot H_t \qquad h_t \overset{not.}{=} K_t y_t^g - d_t$$

Proof. In practical economic applications with real world data, the linear models have to face up to the following constraint: the decision variables and sometimes the state variables have to be constrained in the sense of inequalities and hence the characterization of decision regions is very important. Generally, the decision-maker is restricted in the use of the instruments. Very often, the optimal policies tend to fluctuate with a large amplitude (the stochastic solutions will have a certain dispersion). Instruments variability will generally have a large influence, increasing the error on the targets. The observable fluctuations in instruments are due to the initial impact of the unpredictable shocks and forecast errors. One remedy to avoid drastic changes from one period to another is to impose preselected upper and lower bounds on the values of the control variables (see, SANDBLOM & BANASIK, 1985). These bounds are chosen at the beginning of the application and kept fixed till the end.

Remark 8. However, this remedy may introduce new sources of error. The bounds on instrument variation will produce truncated distributions and so will introduce a bias on instrument variation. Moreover, the variation of instruments is given not by their actual efficiency, but by their relative position in the set of available instruments. Consequently, this will generate discontinuities in the relation between instrument efficiency and optimal policy.

At each period t, the agent will maximize his expected utility function under a set of dynamic constraints imposed in order to avoid drastic changes in the control variable:

$$\arg\max_{x_t \in \Lambda_t} E_{t-1} U_t [W_{[1,t]}, \varphi_t]$$

$$s.t. \quad : \begin{cases} y_t = A_t y_{t-1} + C_t x_t + B_t z_t + D_t + u_t & \\ \underline{L}'_t \leq x_t \leq \overline{L}'_t & \text{(amplitude bounds)} \\ \underline{L}''_t \leq x_t - x_{t-1} \leq \overline{L}''_t & \text{(change bounds)} \\ y_0, y_t > 0, \ \ t = 1, ..., T & \text{(economic constraints)} \end{cases}$$

where

$$W_{[1,t]}(y_1, ..., y_t) \stackrel{def}{=} \sum_{s=\overline{1,t}} W_s(y_s)$$

with W_t an asymmetrical quadratic local loss function, strictly convex and twice differentiable:

$$W_s(y_s) \stackrel{def.}{=} (y_s - y_s^g)' K_s (y_s - y_s^g) + 2(y_s - y_s^g)' d_s = y_s' K_s y_s - 2y_s' h_s + f_s$$

$$h_s \stackrel{not.}{=} K_s y_s^g - d_s \qquad f_s \stackrel{not.}{=} y_s^{g\prime}(h_s - d_s)$$

$$\Lambda_t \stackrel{not.}{=} \left\{ x_t \mid 0 < \underline{L}'_t \leq x_t \leq \overline{L}'_t \text{ and } \underline{L}''_t \leq x_t - x_{t-1} \leq \overline{L}''_t \right\} \subset R^q$$

(the agent's feasible strategies space at time t)

The first set of constraints is imposed in order to keep the instruments within specific positive bounds through time. Constraints of economic rationality impose that $\underline{L}'_t$, $\overline{L}'_t$ and y_t take some positive values. Possible negative realizations of the instruments are ruled out. The wider the bound on the instrument, the higher the importance given by the decision-maker to the variation of the instrument in that direction so that they fit to the active learning process. The amplitude bounds allow the use of more instruments than targets.

It is assumed that the decision-maker chooses these bounds at each iteration of the control algorithm. Thus, he can exploit the information on the previous instruments when fixing the bounds for the next instruments, by allowing a greater variability for an efficient instrument rather than for an inefficient one. The bounds on the instruments are simply the limits up to which the decision-maker decides to extend the research of the optimal solution at each iteration.

As regards the last set of constraints, this indicates that the variation of the control variable between two consecutive periods lies within a prespecified bounded interval. The values of this variation can be either positive or negative.

The two sets of constraints taken together are called boundary conditions. They restrict the set of potential optima. We have the following inequalities:

$$\min \{\underline{L}'_1, ..., \underline{L}'_T\} \leq x_t \leq \max \left\{\overline{L}'_1, ..., \overline{L}'_T\right\} \ \forall \ t = 1, ..., T$$

$$\min \{\underline{L}''_1, ..., \underline{L}''_T\} \leq x_t - x_{t-1} \leq \max \left\{\overline{L}''_1, ..., \overline{L}''_T\right\} \ \forall \ t = 1, ..., T$$

The bounded control approach with bounds not only on the magnitude but also on the rate of change of the controls holds much benefit.

Remark 9. If $\mathbf{\Omega} \subset \mathbf{R}^q$ is the set of admissible values of the instruments from the decision-maker's point of view, then an optimal solution of the above optimization problem which is to be acceptable to the decision-maker will must belong to $\mathbf{\Omega}$. This set is not specified in terms of analytical restrictions, but supposed to exist in the mind of the decision-maker. The fact that $\mathbf{\Omega}$ is not specified is, in practice, almost invariably the case. If the decision-maker can specify $\mathbf{\Omega}$ analytically, these restrictions can be incorporated in the feasible region of the model,

$$\widetilde{\Lambda} \overset{def.}{=} \left\{ x_t \mid \underline{L}_t' \leq x_t \leq \overline{L}_t' \text{ and } \underline{L}_t'' \leq x_t - x_{t-1} \leq \overline{L}_t'' \mid t = 1, ..., T \right\}$$

We note that the nonnegativity constraints on the state and control variables are never binding (dependent each other) in an optimal plan (see, Epstein, 1981).

Remark 10. In the context of a non-linear model, a local linearization allows for the possible bounds on the objectives (whose values depend on the decision vector) to be transformed into linear inequality constraints on the exogenous instruments. Note also that possible bounds on the instruments can be transformed into linear inequality constraints on the objectives.

During the entire period of control, a revision process of the feedback information is required. New information resolves the uncertainty step by step. The value of the optimal instrument $\widehat{x}_t$ is hence obtained by a revision of the expectations at each previous stage of the control. In other words, the agent's optimal decisions evolve over time as result of the periodic learning.

Following the same reasoning employed in **Proposition 1** for an arbitrary period t, we obtain without difficulty the analytical formulas for the feedback optimal equations.

We can write:

$$\widehat{x}_t = \arg\max_{x_t} E_{t-1} U_t(W_{[1,t]}(y_t), \varphi_t) = \arg\max_{x_t} E_{t-1}[\exp\{-(\frac{\varphi_t}{2}) W_t(y_t)\}]$$

where:

$$E_{t-1}[\exp\{-(\frac{\varphi_t}{2}) W_t(y_t)\}] = E_{t-1}[\exp\{-(\frac{\varphi_t}{2})(y_t' H_t y_t - 2y_t' h_t + f_t)\}]$$

with:

$$H_t \overset{def.}{=} K_t, \quad h_t \overset{not.}{=} K_t y_t^g - d_t, \quad f_t \overset{not.}{=} y_t^{g\prime}(h_t - d_t)$$

Substituting $A_t y_{t-1} + C_t x_t + B_t z_t + D_t + u_t$ for y_t, we obtain finally:

$$\overline{V}_t \overset{not.}{=} E_{t-1}[\exp\{-(\frac{\varphi_t}{2}) W_t(y_t)\}] = E_{t-1}[\exp(\omega_2(u_t))] \exp \omega_1(I_{t-1}, x_t, z_t, \beta_t, K_t, d_t, y_t^g)$$

$$= \exp \omega_1(I_{t-1}, x_t, z_t, \beta_t, K_t, d_t, y_t^g) \int_{R^p} (2\pi)^{-\frac{p}{2}} \mid \det \mathbf{\Psi} \mid^{-\frac{1}{2}} \exp\{-\frac{1}{2} \widetilde{u}_t' \mathbf{\Psi}^{-1} \widetilde{u}_t\} \exp \omega_2(\widetilde{u}_t) \, d\widetilde{u}_t$$

with $\beta_t \overset{not.}{=} (A_t, B_t, C_t, D_t)$ and $\omega_2(\widetilde{u}_t)$ a quadratic function in $\widetilde{u}_t$.

We can write:

$$I \overset{not.}{=} \int_{R^p} (2\pi)^{-\frac{p}{2}} \mid \det \mathbf{\Psi} \mid^{-\frac{1}{2}} \exp\{-\frac{1}{2} \widetilde{u}_t' \mathbf{\Psi}^{-1} \widetilde{u}_t\} \exp \omega_2(\widetilde{u}_t) \, d\widetilde{u}_t$$

$$= \mid \det(\mathbf{\Psi}^{-1} + \varphi_t \cdot H_t) \mid^{-\frac{1}{2}} \mid \det \mathbf{\Psi} \mid^{-\frac{1}{2}} \int_{R^p} (2\pi)^{-\frac{p}{2}} \mid \det(\mathbf{\Psi}^{-1} + \varphi_t \cdot H_t) \mid^{\frac{1}{2}} \exp \omega_3(\widetilde{u}_t) \, d\widetilde{u}_t$$

with $\omega_3(\widetilde{u}_t)$ a quadratic function in $\widetilde{u}_t$. Now, we need $\overline{u}_t \in R^p$ such that:

$$\omega_3(\widetilde{u}_t) = -\frac{1}{2}(\widetilde{u}_t - \overline{u}_t)'(\mathbf{\Psi}^{-1} + \varphi_t \cdot H_t)(\widetilde{u}_t - \overline{u}_t) + \text{independent of } \widetilde{u}_t$$

As before, we obtain:

$$\overline{u}_t = -\varphi_t \cdot (\mathbf{\Psi}^{-1} + \varphi_t \cdot H_t)^{-1}[K_t \cdot (A_t y_{t-1} + C_t x_t + B_t z_t + D_t) - h_t]$$

Following the same steps as in **Proposition 1**, the integral becomes:

$$I = \mid \det(I_p + \varphi_t \cdot \mathbf{\Psi} \cdot H_t) \mid^{-\frac{1}{2}} \exp \omega_4(I_{t-1}, x_t, z_t, \beta_t, K_t, d_t, y_t^g)$$

It follows that:

$$\overline{V}_t \overset{not.}{=} E_{t-1}[\exp\{-(\frac{\varphi_t}{2})W_t(y_t)\}] = \exp \omega_1(I_{t-1}, x_t, z_t, \beta_t, K_t, d_t, y_t^g) \cdot I$$

$$= \mid \det(I_p + \varphi_t \cdot \mathbf{\Psi} \cdot H_t) \mid^{-\frac{1}{2}} \exp \omega_5(I_{t-1}, x_t, z_t, \beta_t, K_t, d_t, y_t^g)$$

with:

$$\omega_5(I_{t-1}, x_t, z_t, \beta_t, K_t, d_t, y_t^g) = -(\frac{\varphi_t}{2})[y_{t-1}' A_t' \widetilde{H}_t C_t x_t + x_t' C_t' \widetilde{H}_t (A_t y_{t-1} + B_t z_t + D_t)$$

$$+ x_t' C_t' \widetilde{H}_t C_t x_t + (z_t' B_t' + D_t') \widetilde{H}_t C_t x_t] + \varphi_t x_t' C_t' [I_p - \varphi_t \cdot K_t (\mathbf{\Psi}^{-1} + \varphi_t \cdot H_t)^{-1}] h_t + \text{independent of } x_t$$

where:

$$\widetilde{H}_t \overset{not.}{=} K_t - \varphi_t \cdot H_t M_t^{-1}(\varphi_t) H_t$$

$$M_t(\varphi_t) \overset{not.}{=} \mathbf{\Psi}^{-1} + \varphi_t \cdot H_t = \mathbf{\Psi}^{-1}(\varphi_t \mathbf{\Psi} + H_t^{-1}) H_t$$

The first order condition in x_t writes:

$$-C_t' \widetilde{H}_t A_t y_{t-1} - C_t' \widetilde{H}_t (B_t z_t + D_t) + C_t' [I_p - \varphi_t \cdot K_t (\mathbf{\Psi}^{-1} + \varphi_t \cdot H_t)^{-1}] h_t = C_t' \widetilde{H}_t C_t x_t$$

It follows that:

$$\widehat{x}_t(I_{t-1}, z_t, \beta_t, K_t, d_t, y_t^g) \mid y_0 = \overline{G}_t \cdot y_{t-1} + \overline{g}_t, \qquad t = 1, ..., T$$

where:

$$\overline{G}_t \overset{not.}{=} -(C_t' \overline{H}_t C_t)^{-1}(C_t' \overline{H}_t A_t)$$

$$\overline{g}_t \overset{not.}{=} -(C_t' \overline{H}_t C_t)^{-1} C_t' [\overline{H}_t (B_t z_t + D_t) - (I_p - \varphi_t K_t (\mathbf{\Psi}^{-1} + \varphi_t \cdot H_t)^{-1}) h_t]$$

$$\overline{H}_t \overset{not.}{=} K_t - \varphi_t \cdot H_t M_t^{-1}(\varphi_t) H_t$$

The coefficients of the optimal policy (behavioral equation) depends, in this new formulation of the problem, on the risk-aversion index specific to each period of control. This is expected to contribute to the equilibrium and stability of the system as well as to improve the agent's future utilities. The optimal feedback control will stabilize the system because it allows for a free flow of information about the system evolution.

The decision rule is systematically reviewed and revised in response to new signals from the environment. The decision-maker will thus refine the distance between the current target variable and the fixed system characteristics. The deviations from the targets will be minimal amongst all possible deviations, because the imperfections on $\widehat{x}_1, ..., \widehat{x}_{t-1}$ will not affect $\widehat{x}_t$. Thus, the optimal policy is robust to the variance of the shocks. It is interesting to note that the sufficient variables for describing $\widehat{x}_t$ belong to some spaces of constant dimension, while the endogenous information set I_t generates a sequence of spaces of increasing dimension.

The parameters of the behavioral equation are related to the parameters of both economic environment and objective function. The former are derived from the latter through optimization. By consequence, if the parameters of the economic environment (or of the objective function) change, the parameters of the behavioral equation will also change. Knowledge of the former parameters can be used to derive the parameters of the behavioral equation, which can then be utilized to obtain forecasts of the endogenous target variable. The parameters of the state equation also change if the generating mechanism for x_t changes.

The existence of the optimum may be thus restricted to certain configurations of the parameters of interest. Accurate estimates are necessary for an efficient implementation of the above numerical methods. They represent a basic information for an optimal algorithm and hence are very important as regards the accuracy of the numerical simulations.

Consider first the case when there exist an optimum for each period of control. We give in this sense a graphical representation of the agent's optimal actions during the entire planning horizon as well as the simulated performance of the system with respect to the fixed targets.

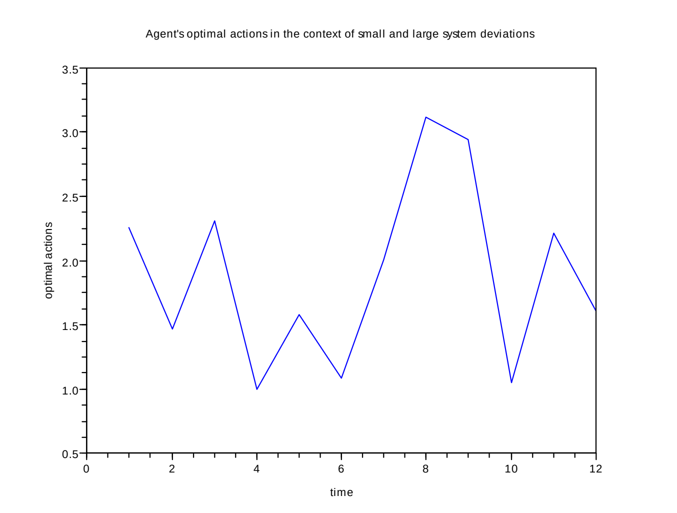

Figure 1:

For this scenario, it is supposed that the agent's objective is to keep the instruments within the optimal interval from 0 to 3.5

When there is no optimal solution at a given period of time, the decision-maker will choose the most recent solution which has conducted the system close to its optimal target. Three distinct scenarios can be imagined here:

i) the variable of control takes a negative value. We give below a suggestive graphic in this sense.

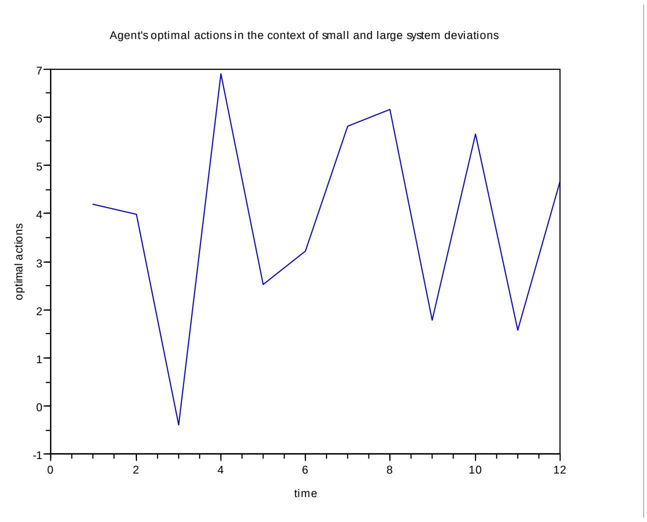

Figure 2:

For this scenario, the agent obtains a negative value for $\widehat{x}_3$. In this case, he will choose for the period $t = 3$ between $\widehat{x}_1$ and $\widehat{x}_2$, depending on their performance with respect to the targets y_1^g and respective y_2^g.

ii) the variable of control does not satisfy the condition regarding the amplitude bounds. For an illustration, we give below a suggestive graphic.

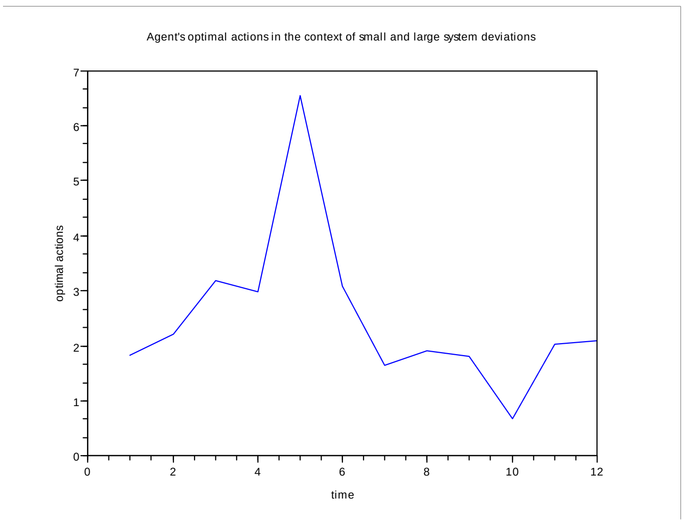

Figure 3:

For the above scenario, the agent's objective is to keep the instruments within the interval of interest $(0, 4)$ during the control period. For the period $t = 5$, the condition of amplitude limitation is not satisfied while the change bounds (-4 and 4) are not exceeded.

iii) the variable of control does not satisfy the condition regarding the change bounds.

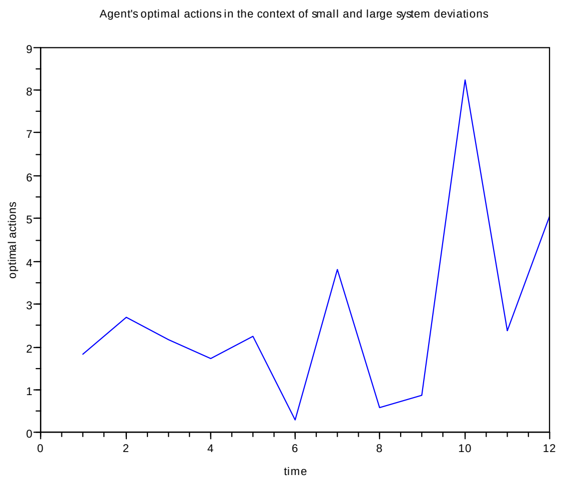

Figure 4:

The agent's objective, for this scenario, is to keep the instruments within the interval $(0, 5]$. In this case, both boundary conditions are not satisfied for the period $t = 10$.

It is important to note that is possible to implement the same optimal action for distinct periods of time. We illustrate this possibility by a suggestive graphic.

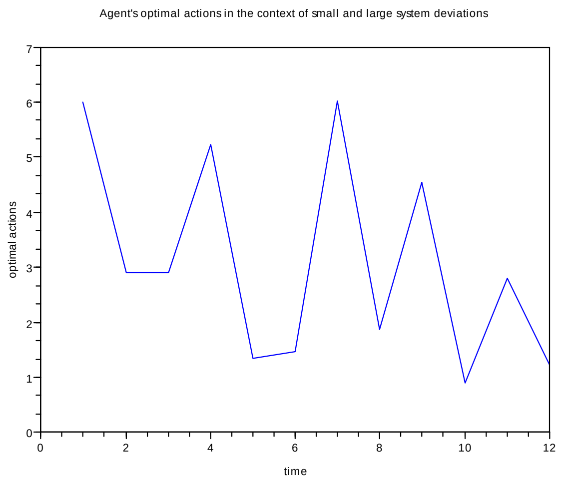

Figure 5:

The agent's objective is to obtain small deviations of the system with respect to the fixed targets during the entire control period. We give below a suggestive graphical representation for this scenario.

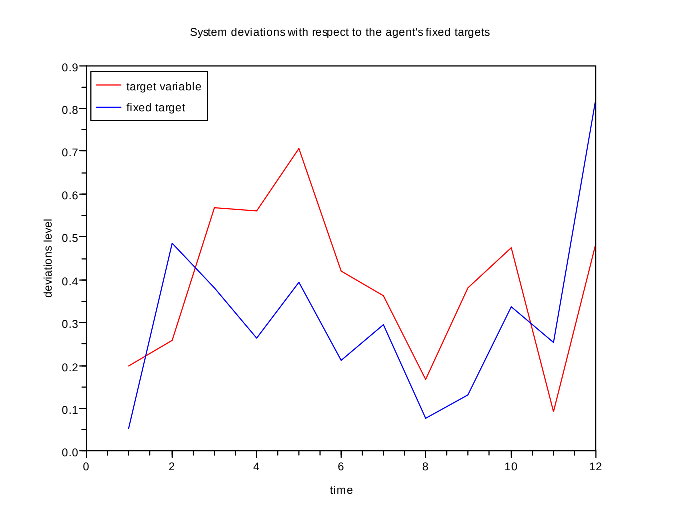

Figure 6:

However, in the case of a very noisy environment, severe problems can be caused to the agent in optimizing the system trajectory. Two distinct scenarios can be considered to illustrate this possibility:

1) the agent does not succeed to constrain the system to follow the optimal trajectory η during the entire planning horizon.

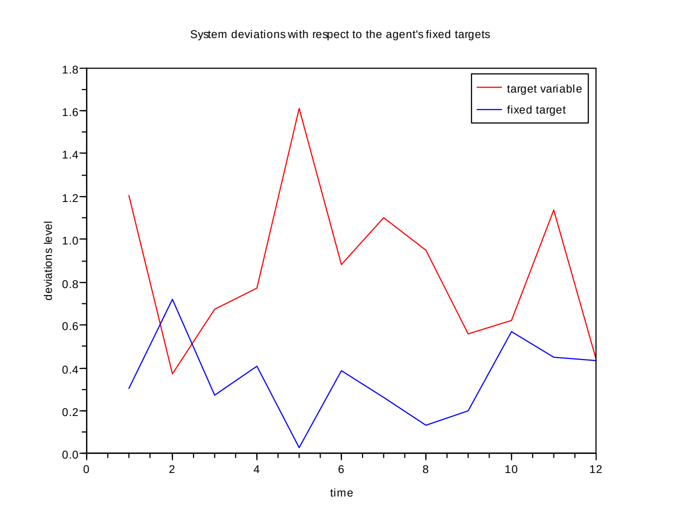

Figure 7:

2) the agent exceeds considerably the fixed optimal targets at each period of the planning horizon. All the deviations of the system are (very) large, that is:

$$\parallel y_t - y_t^g \parallel \gg 1, \qquad \forall\ t = 1, ..., T.$$

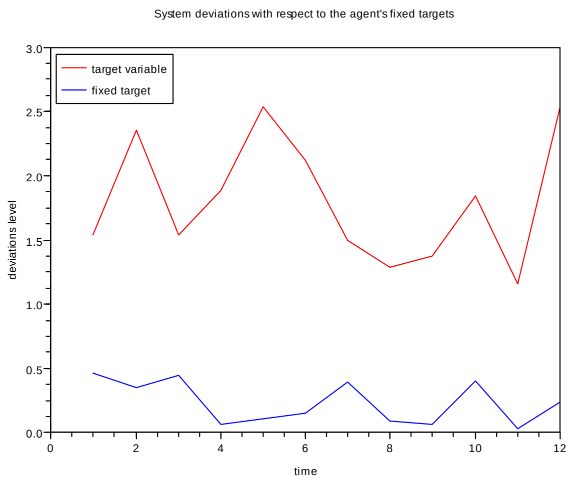

Figure 8:

We present now a short analysis of the agent's optimal actions with respect to the risk-aversion index level during the control period. Three distinct cases are discussed here: **i**) when the agent is chracterized by a small risk-aversion; **ii**) when the agent is characterized by a high risk-aversion; **iii**) when the agent's risk-aversion index is fluctuating between -1 and 0. We give below two superposed graphics that illustrate the first two cases.

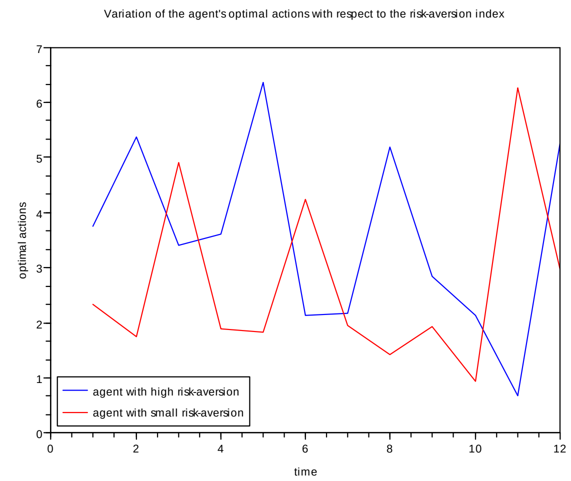

Figure 9:

The optimal actions are generally different for different attitudes to risk during the planning horizon. They are not monotone functions with respect to the risk-aversion index parameter. The dependence between $\widehat{x}_t$ and φ_t is non-linear. The agent's optimal actions can be smaller or higher in magnitude, depending on the context of the problem. They are not necessarily correlated with the agent's risk-aversion type.

We illustrate below the case where the risk-aversion index is fluctuating between -1 and 0. An exceeding of the threshold $\varphi_{\min}$ will have a non-negligible local effect on the agent's optimal policy.

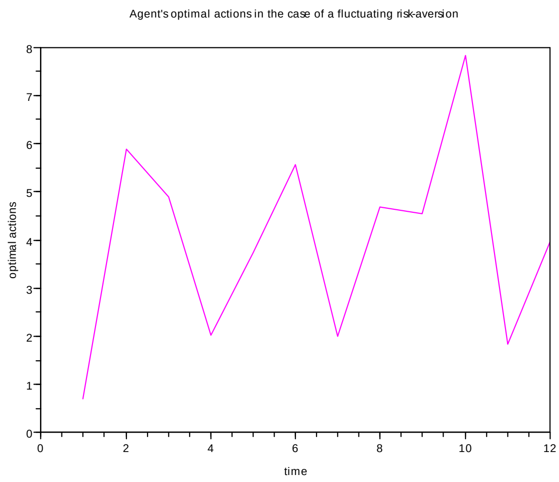

Figure 10:

It is crucial for the decision-maker to optimally choose the weighting parameters K_t and d_t before proceeding at the maximization of his objective function on the regulation horizon.

In general, K_t is chosen to be a symmetrical positive semi-definite diagonal matrix attaching penalty constant weights to deviations of the state variable from its desired level. If K_t is not diagonal, then penalties also attach to covariances of deviations of the state variable from its desired level.

The variations of the weighting parameters will affect the extensiveness of the agent's loss function. It is therefore very important to know the effect of K_t and d_t on the agent's behavior during the process of optimization and control.

The role played by the ponderation matrix K_t can be easily illustrated in the univariate model case. Thus, using the matrix differential rules, one obtains the first-order condition for $\triangle\, y_t$:

$$2K_t \triangle y_t + 2d_t = 0 \Leftrightarrow \quad \triangle y_t = -\frac{d_t}{K_t} \text{ (if } K_t \neq 0)$$

By consequence, if $d_t \neq 0$, each increase (decrease) of the parameter K_t causes an increase (decrease) of the distance between the value of y_t measured and that one fixed at time t. We illustrate this behavior for the two cases discussed above.

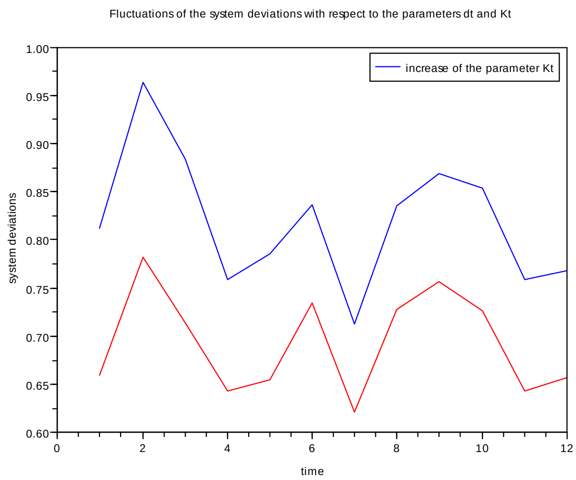

Figure 11:

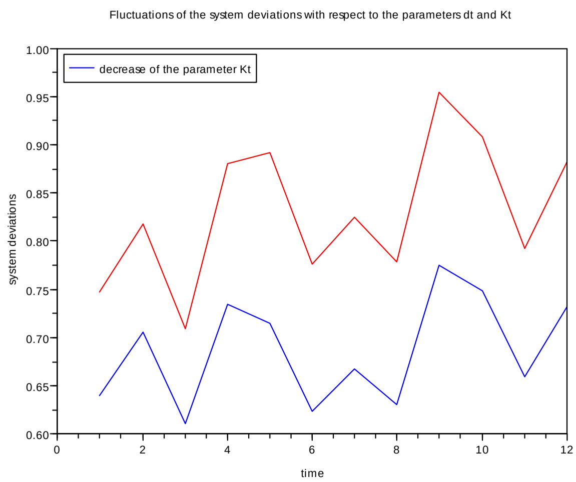

Figure 12:

The learning algorithm presented in this paper solves completely the linear-quadratic control problem with suitable initial and boundary conditions in the case where the decision-maker is characterized by an endogenous risk-aversion during the entire planning horizon. A non-negligible advantage of this strategic decision rule is that it is more simple to compute and easy to use. This new control approach has the potential to better predict the behavior of the system over time because the learning algorithm implemented by the decision-maker integrates a controlled risk-aversion at each step of the control. It improves the decision-maker's ability to understand the environment response to some of his actions. Moreover, in the context of a dynamic stochastic game, it will improve the agent's ability to understand a rival's pattern of play.

8. Concluding Remarks and Possible Extensions

The purpose of this paper is to correct and improve the formulas for the feedback linear optimal equations obtained by Van der Ploeg (1984a, 1984b) and Chow (1973, 1976a, 1976B, 1977, 1978, 1981, 1993) in the classic context of a controlled decision process, when the decision-maker is characterized by a constant exogenous risk-aversion.

We extend the previous studies to the more realistic and attractive case of a controlled endogenous risk-aversion. The present paper offer to decision-makers (e.g., governments, firms, economic agents) decision rules that allow for a better management and control of dynamic stochastic environments characterized by important endogenous fluctuations.

We present here possible directions for future research, including both theoretical extensions and empiric work.

We can investigate, from this new perspective, the case of a total separation between the control of the process and the control of the instrument. In this particular context, the optimal control will be independent of the accuracy of the information on the current state. The separation of parameter estimation from the decision making process generally yields lower utility than an integrated approach which takes account of estimator uncertainty. The resulting utility loss can be substantial. It is expected that the precision of the combined estimate is greater than the precision of separate estimates, because it allows for a better evaluation of the true. The problem can be extended easily to the case when the parameters of the model satisfy some equality restrictions. Such restrictions usually relate to a small class of control instruments and do not necessarily hold for all equations.

An other important application is the case of an asymmetric criterion function expressed as a sum of weighted squares of deviations from given target values for the objectives and instruments. The desired values of the instruments are thus included in the quadratic loss function in order to prevent them from going too far away from realistic values. However, we must be conscious that no unique criterion unambiguously determines the target values for the instruments. In this formulation of the problem, it is taken into account the importance given by the decision-maker to the reduction of the difference between the instruments and their corresponding target values. In other words, the corresponding weighting parameters are related to the relative loss attributed to the non-achievement of the target values for the instruments.

The case of endogenous targets is an other good topic for further research. The fixed goal is flexible with respect to the possible changes (the nature can change its goal) and can be modified (in a new formulation of the problem) without incurring additional cost, time, or effort. A decision problem is often redefined during the decision process itself. This is because, usually, a target path is prescribed without any consideration of the question whether it can be obtained. In this more general context, it is expected that the agent's attitude to risk will be improved and thus, the implemented actions will be more consistent with the planned objectives.

Economic intuition tell us that the learning may produce stabilizing and destabilizing effects. Exogenous stochastic shocks can perturb the system being propagated forward by the learning rules. When the decision-maker is uncertain about the dynamic of the system, the process of learning is generally expected to generate local instability. There are potential effects of a change in the parameters of the system on the stable steady states of the system. The estimation of the parameters of interest may be a possible source of instability. In other words, there exists a trade-off between learning and instability. Usually, the data are not rich enough to estimate successfully the structural models. After computing the optimal policy, the decision-maker's objective is to analyze the dynamic behavior of the controlled stochastic process. Is it stationary or explosive? In other words, are the target variables stable or not? Do distributed lags have any effect on system stability? If the system is stable but uncontrollable, then the state variable remains bounded and almost unaffected by the choice of the control variable. This type of analysis is necessary in most practical control applications. In the literature on control theory, only little effort has been spent on stability issues in stochastic environments (see, Chow, 1993). Because the system is stochastic, the meaning of stability is ambiguous. It highlights the need for further research in this exciting area of research.

Although linear models are conceptually convenient and analytically flexible, they do not always provide an adequate framework for modelling the economic behavior. More complex models are needed if policy intervention is the purpose of modelling. A comparative empirical study between several alternative approaches (e.g., non-linear modelling, bayesian learning, semiparametric /nonparametric model specification) is a good topic for further research. We need empirical knowledge of the optimal policy performance in different contexts of decision making, from this new risk-averse perspective. Dynamic numerical simulations may be used to obtain information on how economic agent's optimal decisions vary over time. Generally, different methods may lead to important differencies when these are used on simulated data and to different statistics. Different contexts call for different actions. An interesting question is whether the differences between two distinct procedures lead to qualitatively different predictions on the form that risk-averse learning should take. This problem deserves special attention. Work remains to be done for more general specifications.

The analysis can be extended to the case of imperfect ex-post information. A study of economic effects of information and uncertainty is very useful. In the literature on control, it is supposed that all the variables are observable, but generally the dynamic information is limited by sample size and masked by noise. This complicates considerably the empiric analysis. There are substantial losses in efficiency when only the subset of data that has complete observations is used in estimation. The quality and the quantity of information is critically important to the decision-maker's learning process. It generally affects the agent's optimal decisions over time. When the sampling process does not identify the distribution of interest (e.g., left-truncated and /or right-censored data), it is very important to use all the available reliable a priori information to analyze the model. This information can be numerical, may take the form of a belief or can be a particular specification suggested by the economic theory. It is crucial to establish in this case what information is observable and what is unobservable. In this case, the classic methods of estimation and optimal decision making will be biased. We need to analyze how change the classic methods in order to ensure the convergence and the bias reduction when allowing for endogenous uncertainty and dynamic risk-aversion.

Of a great interest is also the case of an working horizon which extends as time evolves. The moving horizon length will be thus an endogenous parameter. It allows to incorporate new information of the system at any point of time. It is obtained, in this case, a moving horizon decision rule based on a continuous refinement process of the risk-aversion index. This type of analysis allows to combine the finite and infinite horizon optimization problems in the context of a controlled endogenous risk-aversion.

Significant differences exist between an individual control problem (viewed as a game against nature), when the agent is submitted only to environment constraints, and respective a controlled dynamic game (where each player is, in addition, constrained by the opponent's behavior). Many interesting applications of real interest in the context of dynamic stochastic games (Nash or Stackelberg) can be exploited from this new perspective. In this case, the equilibrium of the game is subjected to many constraints which mix the parameters of interest. Two distinct cases can be considered, depending upon the nature of the game (the type of interactions): cooperative or non-cooperative behavior. The objective, in each case, is the same: obtaining the system stability with optimal risk-sharing between the players. We can also test (under heteroskedasticity) if a given discrete time-series arises from a game with closed-loop Nash / Stackelberg strategies. In the literature, the preocupation on the (theoretical and empirical) tests aspects in controlled dynamic stochastic games is almost unexistent. We encourage other researchers to take up the challenge.

References

[**1**] AMEMIYA, T., 1974. Non-Linear Regression Models, Chapter **6**, **333-389**.

[**2**] AMMAN, H. M., 1995. Numerical methods for linear-quadratic models. In: Amman, H. M., Kendrick, D. A., Rust, J. (Eds.), Handbook of Computational Economics. North-Holland, Amsterdam.

[**3**] AMMAN, H. AND KENDRICK, D., 1995. "Nonconvexities in stochastic control models". *International Economic Review* **36** (2), **455-475**.

[**4**] ANDERSON, B. AND J. MORE, 1979. Optimal filtering (Prentice-Hall, Englewood Cliffs, N J).

[**5**] ANDREWS, D. W. K., 1991. "Heteroscedasticity and Autocorrelation Consistent Covariance Matrix Estimation", *Econometrica*, **59**, **817-858**.

[**6**] ARROW, K. J., 1971A. The Theory of Risk Aversion, Amsterdam: North-Holland.

[**7**] ARROW, K. J., 1971B. Essays in the Theory of Risk Bearing, Chicago: Markham Publishing Co.

[**8**] BALVERS, R., AND COSIMANO, T., 1990. "Actively learning about demand and the dynamics of price adjustment". *The Economic Journal,* **100**, **882-898**.

[**9**] BARBOSA, F., 1975. "Rational Random behaviour, Doctoral dissertation". University of Chicago Press.

[**10**] BECK, G. W AND V. WIELAND, 2002. "Learning and Control in a Changing Economic Environment". *Journal of Economic Dynamics & Control*, **26**, **1359-137**7.

[**11**] BELLMAN, R., 1961. Adaptive Control Processes: A Guided Tour. Princetown University Press.

[**12**] BOCK, V. W. G. AND P. PAULY, 1978. "An optimization approach in multiple target problems using inequality constraints: The case against weighted criterion function". *Annals of Economic and Social Measurement,* **6,** 613-630.

[**13**] CHOW, G. C., 1973. "Effect of Uncertainty on Optimal Control Policies", *International Economic Review*, **14**, **632-645**.

[**14**] CHOW, G., 1976A. Analysis and Control of Dynamic Economic Systems, John Wiley & Sons, New York.

[**15**] CHOW, G., 1976B. "The Control of Nonlinear Econometric Systems with Unknown Parameters", *Econometrica*, **44**, **685-69**5.

[**16**] CHOW, G., 1977. "Usefulness of Imperfect Models for the Formulation of Stabilization Policies", *A.E.S.M.*, **6**, **175-187**.

[**17**] CHOW, G., 1978. "Evaluation of Macroeconomic Policies by Stochastic Control Techniques", *International Economic Review*, **19**, **311-320**.

[**18**] CHOW, G., 1981. Econometric Analysis by Control Methods, John Wiley & Sons, New York.

[**19**] CHOW, G., 1993. "Optimal control without solving the Bellman equation", *Journal of Economic Dynamics & Control* **17**, **621-63**0.

[**20**] CRUZ, J. B., 1975. "Survey of Nash and Stackelberg Equilibrium Strategies in Dynamics Games". *Annals of Economic and Social Measurement*, **4, 2**: **339-344**.

[**21**] DAVIDSON, R. AND J. G. MACKINNON, 1981. "Several tests for model specification in the presence of alternative hypotheses", *Econometrica*, **3**, **781-793**.

[**22**] EASLEY, D. AND N. KIEFER, 1988. "Controlling a stochastic process with unknown parameters". *Econometrica,* **56**, **1045-106**4.

[**23**] ENGLE, R. F., 1974. "Specification of the disturbance for efficient estimation", *Econometrica*, **42**, **135-146**.

[**24**] EPSTEIN , L. G., 1981. "Duality theory and functional forms for dynamic factor demands". *Review of Economic Studies*, **48**, **81-96**.

[**25**] FLEMING, W. H. AND R. W. RISHEL, 1975. Deterministic and Stochastic Optimal Control (Springer-Verlag, Berlin).

[**26**] GOURIEROUX, C. AND A. MONFORT, 1995. "Testing, encompassing and simulating dynamics econometric models", *Econometric Theory*, **11**, **195-228**.

[**27**] GOZALO, P. AND O. B. LINTON, 1995. "Using Parametric Information in Nonparametric Regression", Working paper **95-40**, Brown University.

[**28**] GRANGER, J. AND TERASVIRTA, T., 1993. "Modelling Non-linear Economic Relationships". Oxford: *Oxford University Press*.

[**29**] HAAVELMO, T., 1944. "The Probability Approach in Econometrics", Econometrica, **12**, **1-118**. *Supplement*.

[**30**] HARVEY, A. C., 1990. "Forecasting, Structural Time Series Models and the Kalman Filter". *Cambridge University Press*.

[**31**] HECKMAN, J. J., 1979. "Sample bias as a Specification error", *Econometrica*, **1**, **153-161**.

[**32**] HENDRY, D. F., 1995. "Dynamic Econometrics". Oxford: *Oxford University Press*.

[**33**] JACOBSON, D. H., 1973. "Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games", Institute of Electric and Electronics Engineers Transactions on Automatic Control AC-**18**, 124-131.

[**34**] JACOBSON, D. H., 1977. "Extensions of Linear-Quadratic Control, Optimization and Matrix Theory". London, Academic Press.

[**35**] JUDD, K., 1998. "Numerical Methods in Economics". *MIT Press*, Cambridge, MA.

[**36**] KALMAN, R. E., 1960. "A New Approach to Linear Filtering and Prediction Problems". *Journal of Basic Engineering*, **82**, **35-45**.

[**37**] KALMAN, R. E. AND R. S., BUCY, 1971. "New results in Linear Filtering and Prediction Theory". *Journal of Basic Engineering*, **83**, **95-107**.

[**38**] KARATZAS, A. AND S. E. SHREVE, 1988. Brownian motion and stochastic calculus (Springer-Verlag, Berlin).

[**39**] KARP, L., 1987. "Risk-Aversion in a Dynamic Trading Game", WP. **404**, Division of Agriculture and natural Resources, University of California.

[**40**] KENDRICK, D. A., 1981. Stochastic control for economic models. Economic handbook Series. McGraw-Hill, New York.

[**41**] KIEFER, N. AND Y. NYARKO, 1989. "Optimal control of an unknown linear process with learning". *International Economic Review,* **30**, **571-58**6.

[**42**] KLEIN, R. W., RAFSKY, L. C., SIBLEY, D. S., AND WILLING, R. D., 1978. "Decisions with estimation uncertainty". *Econometrica*, **6**, **1363-1387**.

[**43**] KNIGHT, F. H., 1971. "Risk, uncertainty and profit". University of Chicago press, Chicago, IL.

[**44**] KOYCK, L. M., 1954. Distributed lags and Investment Analysis. Amsterdam: North Holland. Publishing Company.

[**45**] LELAND, J. W., 1990. A Theory of Approximate Expected Utility Maximization. Mimeo, Department of Social and Decision Sciences, Carnegie Mellon University.

[**46**] MANUEL, S. SANTOS AND JESÚS VIGO-AGUIAR, 1998. "Analysis of a numerical dynamic programming algorithm applied to economic models". *Econometrica*, **2**, **409-426**.

[**47**] MARCET, A., 1994. "Simulation Analysis of Dynamic Stochastic Models: Application to Theory and Estimation", in Advances in Econometrics, Sixth World Congress. Vol 2,

ed. by C. Sims. Cambridge: Cambridge Univ. Press.

[**48**] NEWEY, W. K. AND K. D. WEST, 1987. "A Simple Positive Semi-definite Heteroscedasticity and Autocorrelation Consistent Covariance Matrix", *Econometrica*, **55**, **703-708**.

[**49**] PITCHFORD, J. D., 1977. Two state variable problems. In: Pichford, J. D., Turnovsky, S. J., (Eds.), Applications of Control Theory to Economic Analysis, Contributions to Economic Analysis, vol. **101**, North-Holland, New York.

[**50**] PRATT, J. W., 1964. "Risk Aversion in the small and in the large", *Econometrica*, **32**, **122-136**.

[**51**] ROSENBERG, B., 1968. "Varying parameter estimation". Unpublished Ph.D. dissertation. Department of Economics, Harvard University.

[**52**] ROSS, S. A., 1981. "Some stronger measures of risk aversion in the small and in the large with applications", *Econometrica*, **3**, **621-638**.

[**53**] PROTOPOPESCU, D., 2007. "Improving the Risk Concept: A Revision of Arrow-Pratt Theory in the Context of Controlled Dynamic Stochastic Environments",Working paper **No. 727.08**, CODE, Universitat Autònoma de Barcelona, Spain.

[**54**] RUST, J., 1996. Numerical Dynamic Programming in Economics, in Handbook of Computational Economics, ed. by K. M. Amman, D. A. Kendrick, and J. Rust. Amsterdam: Elsevier Science.

[**55**] RUSTEM, B. AND M. B., ZARROP, 1979. "Optimal Control for Econometric Models: an Approach to Economic Policy Formulation". London, MacMillan.

[**56**] SANDBLOM, C. L. AND J. L. BANASIK, 1985. "Economic policy with bounded controls". *Economic Modelling,* **2**, **135-148**.

[**57**] SARGAN, D., 1988. Lectures on Advanced Econometric Theory, Oxford, Basil Blackwell.

[**58**] SARGENT, T. J., 1987. Macroeconomic Theory. Academic Press, Orlando.

[**59**] TAYLOR, J. B. AND H. UHLING, 1990. "Solving Non-Linear Stochastic Growth Models: A Comparation of Alternative Solution Methods", *Journal of Business and Economic Statistics*, **8**, **1-18**.

[**60**] VAN DER PLOEG, L., 1984A. "Risk and Linear-Quadratic Stabilization", *Economic Letters*, **15**, **73-78**.

[**61**] VAN DER PLOEG, L., 1984B. "Economic Policy Rules for Risk-Sensitive Decision Making", *Zeitschrift für Nationalöekonomie*, **44, 207-235**.

[**62**] VAN DE STADT, H., KAPTEYN, A., VAN DE GEER, S., 1985. "The Relativity of Utility: Evidence from Panel Data", *Review of Economics and Statistics*, **2**, **179-187**.

[**63**] WALSH, V., 1996. "Rationality, Allocation and Reproduction". *Clarendon Press,* Oxford.

[**64**] VARIAN, H. L., 1990. "Goodness-of-fit in optimizing models", *Journal of Econometrics,* **46**, **125-140**, North-Holland.

[**65**] WHITTLE, P., 1981. "Risk-sensitive linear-quadratic control problems", *Advances in Applied Probability* **13**, **764-777**.

[**66**] WHITTLE, P., 1989. "Entropy Minimizing and Risk-Sensitive Control Rules", *System and Control Letters*, **13**, **1-7**.

[**67**] WHITTLE, P., 1990. "Risk Sensitive Optimal Control", Cambridge, *Cambridge University Press.*

[**68**] WIEDMER, T., FICHER, A. C. & NAHRAIN, U., 1996. "Pollution Control in a Stochastic Environment". *Swiss Journal of Economics and Statistics.* **132, 4**: **575-590**.

[**69**] WIELAND, V., 2000. "Learning by doing and the value of optimal experimentation". *Journal of Economic Dynamics & Control*, **24**, **501-534**.

[**70**] XEPAPADEAS, A. P., 1992. "Environmental Policy Design and Dynamic Non-point Source Pollution". *Journal of Environmental Economics and Management*, **23**, **1**: **22-39**.