Seventh Annual Pacific-Rim Real Estate Society Conference
Adelaide South Australia, 21st to 24th January 2001

EXTENDING REAL OPTIONS THEORY TO ACCOUNT FOR PROPERTY INVESTMENT UNDER CONDITIONS OF UNCERTAINTY AND INCOMPLETE MARKETS

James Juniper

Lecturer

School of International Business

Contact author for all inquiries
Phone: 8302 0743 (International +61 8302 0743)
Facsimile: +61 (0)8 8302 0512
E-mail: james.juniper@unisa.edu.au

Keywords: Real Options, Risk-sensitive, Control Theory, Incomplete Markets, Property Cycle, Relative Entropy, H-infinity Control.

Abstract: The literature on real options theory is both diverse and rapidly expanding. Now standard techniques for the pricing of derivatives given a stochastic process for the price of underlying assets, are increasingly applied to the case of real investment in productive assets, including property development (Trigeorgis, 1996). Typically, this latter class of investments is represented as a set of options over deferment of a project, termination and salvage, switching of inputs and/or outputs, spawning of related projects, and the expansion, contraction, and temporary shut-down of projects. Concurrently, finance theorists have drawn on the mathematical literature on risk-sensitive and robust control theory under norm bounds and relative entropy constraints as one vehicle for accommodating uncertainty and market incompleteness (Andersen, Hansen, and Sargent 1999; McEneaney, W.M., 1997; Tornell, A., 2000).

This paper provides a heuristic understanding of the relationship between these two bodies of research by examining recent work that establishes range bounds over option prices in incomplete markets (eg, that arising due to the stochastic volatility of stock prices or the existence of a stochastic interest rate) through the application of "good-deal" bounds over the Sharpe ratios and gain-loss ratios of basis assets (Bernardo and Ledoit, 2000; Cochrane and Saa-Requejo, 2000). By varying the stipulated bound, the valuer can move along a spectrum ranging from the set of non-arbitrage bounds through to the uniquely defined option price that is associated with the pricing kernel of a chosen asset pricing model. The "good-deal" bound can thus be interpreted as a measure of investor uncertainty relative to a reference probability distribution for the equilibrium asset-pricing model. The paper identifies the precise relationship holding between the sup-norm bound on the pricing kernel, minimum cross entropy (Stutzer, 1995), and the stochastic uncertainty constraint that is adopted in certain robust control problems. In addition, it examines the derivation of martingale measures and the role of entropy techniques in Generalized Method of Moments estimation. As such, it sets out an agenda for future research into real options-based valuation of investment under uncertainty.

Table of Contents:

1.0 Introduction ____ 4

2.0 Real Options, Financial Options and Property Development ____ 4

3.0 Finance Applications of Risk-sensitive and Robust Control ____ 8

4.0 Option Pricing and General Equilibrium Asset Pricing under Complete Markets ____ 15

5.0 Option Pricing and General Equilibirium Asset Pricing under Incomplete Markets ____ 16

6.0 Stutzer's Entropy-Based Analysis of Asset-Pricing Models ____ 20

7.0 Kitamura and Stutzer's Entropy-based GMM Framework ____ 24

8.0 Minimum Cross-Entropy and Martingale Measures: the Discrete-Range Case ____ 28

9.0 Conclusion ____ 32

Bibliography ____ 35

Technical Appendix: ____ 37

1.0 Introduction

In this paper I selectively review the literature on real options theory and property development. I then examine financial applications of risk-sensitive and robust control theory. I attempt to provide a heuristic bridge between these two fields of study by examining related research on the use of good-deal bounds over Sharpe rations and gain-loss ratios, and minimum cross entropy methods for pricing options in incomplete markets. In regard to the latter, I examine new entropy-based Generalized Method of Moment techniques that appear to have improved small sample properties. Finally, I review Gzyl's discrete-range "maxentropic" derivation of martingales that can be used in option pricing models that are based on the binomial lattice, Markov-chain and finite-difference methodologies. At the same time, I draw on Shore and Johnson's (1980) research based on desirable properties of statistical inference to which entropy measures seem to conform. I observe that Shore and Johnson's analysis relates both to equality and inequality constraints: the latter providing a direct link to approaches based on good-deal bounds and gain-loss ratios.

To the best of my knowledge, no other authors have identified the relations holding between real options theory and this latter strand of inquiry. Nor have the links between this literature and finance-based applications of risk-sensitive stochastic control theory been clearly articulated. Although this is principally a review paper, I hope to identify practical opportunities for future research and quantitative analysis. My ultimate intention, however, is to identify elements which are common to both the risk-sensitive control and option pricing literature to better explain the observed volatility of investment over the property cycle. Alan Greenspan's comments about episodes of "irrational exhuberance" and "uncertainty aversion" in equity markets were equally applicable to the "boom" and "crash" characteristics of the global property cycle over the late 80s (and more recently amongst the Asian economies). It is these aspects of the property cycle are what I am trying to grasp, albeit, in largely theoretical terms.

2.0 Real Options, Financial Options and Property Development

One way to model (real) investment under increasing *risk* is to exploit the analogy between financial and real options (Dixit and Pyndyck, 1994). The following simple correspondences hold between the elements that determine the value of financial (call) options and those that determine the value of real investment projects considered as options over the present value of expected cash flows over the project horizon (Trigeorgis, 1996):

Call	**Real Option**
current value of stock	gross present value of expected cash flows
exercise price	investment cost
time to expiration	window of opportunity
stock value uncertainty	project value uncertainty
riskless interest rate	riskless interest rate

These real options theories of investment, initially developed from the perpetual options models of Paul Samuelson and Robert Merton, have now secured a prominent position in both the theoretical and empirical literature on corporate investment decisions (for theoretical examples see Dixit and Pyndyck 1994, Abel *et al*. 1996, and Guiso and Parigi 1996; and for influential empirical studies see Price 1996, Leahy and Whited, 1996, and Hurn and Wright, 1994). To a large extent this is because the managerial flexibility to revise current or adapt later decisions cannot be captured by conventional discounted cash-flow techniques. Although some aspects of choice and flexibility can be accommodated by combining net present value techniques with decision-tree analysis, Trigeorgis (1996, Chapter 5) shows theoretically that, to properly account for the sequential exercise of real options, discount rates would have to be continually modified as each decision is taken.

Dixit and Pyndyck (1994, Chapter 11) develop a simple model of irreversible investment in which a representative firm faces uncertain demand given by $P = YD(Q)$, where Q equals output, P equals price, and Y is a shift variable following geometric Brownian motion

$$dY = aYdt + sYdz$$

The solution is characterized by the fact that a critical threshold value for Y must be exceeded for investment to take place[1]. The threshold is determined as $Y(K) = \frac{b_1}{b_1 - 1}\frac{dk}{H'(K)}$, where k is the unit

[1] In their model (Dixit and Pyndyck 1994; also see Price 1996, for an empirical application), the firm faces uncertain demand given by $P = YD(Q)$ where Q = output, P = price and Y is a shift variable with geometric Brownian motion $dY = aYdt + sYdz$. Given a production function $G(K)$, the firm's profit can be expressed as: $p = yD[G(K)]G(K) = YH(K)$ where $H''(K) < 0$. Setting up the problem as a dynamic program, the

price of capital. The term $\frac{dk}{\mathrm{H}'(\mathrm{K})}$ reflects the cost of installing an extra unit of capital relative to the expected increase in the present value of the firm. The ratio $\frac{b_1}{b_1 - 1}$ gives the multiple by which the expected present value must exceed the marginal capital cost. In the Cobb-Douglas case, $H(K) = K^q$ and the long-run growth rate of capital can be shown to equal:

$\frac{1}{2}s^2 \frac{\left(\frac{2a}{s^2} - 1\right)}{1-q} = \frac{a - \frac{1}{2}s^2}{1-q}$, so that *increasing* risk *reduces* capital accumulation.

Real options theory can readily be extended beyond the simple option to defer an investment project, to incorporate option values associated with a contraction in the scale of an existing project, an expansion at a particular stage in the life of a project as new information becomes avaiable, the termination and salvage of a project, default on future installments, and switching between various resource inputs or between project outputs[2]. For example, investments in advanced manufacturing technology are frequently a source of switching option values (Lei et. al., 1996).

For obvious reasons, real options theories have increasingly been applied to the study of land and property development. Not only are developers able to make varied decisions about over the timing of their investments, but also decisions about the character of projects (eg. the quality of office accommodation), the density (eg. number of rental units per hectare), and the nature of usage (eg. the relative proportion of commercial and residential units).

---

authors derive the following expression for the value function: $W(K,Y) = YH(K)dt + e^{-rdt}E(W(K,Y+dY))$, which satisfies the second-order differential equation

$\frac{1}{2}s^2Y^2W_{YY}(K,Y) + aYW_Y(K,Y) - rW(K,Y) + YH(K) = 0$. The solution to this equation is given by:

$W(K,Y) = B_1(K)Y^{b_1} + YH(K)/d$ where $d = r - a$ and b_1 is the positive root of

$f = \frac{1}{2}s^2 b(b-1) + ab - r = 0$

Here, $YH(K)/d$ is the expected value of profits the firm would receive if it maintained a constant value of K, while $B_1(K)Y^{b1}$ is the current value of its future option to expand capacity. The constant of integration B_1 can be solved using the ‘value-matching’ and ‘smooth-pasting’ conditions.

2 On a tangential note, real options theories have also been deployed to attack the current obsession with shareholder value-added benchmarks and incentive schemes. These fashionable metrics are largely based on discounted residual income measures of project value: an approach that *completely* ignores the often sizable option multiples that are embodied in project worth.

Titman (1985) has applied real options theory to the choice between construction of six or nine unit apartments. In Capozza and Hedley's model (1990) while the density of development was assumed to be fixed the developer could choose the optimal time for conversion. In contrast, Clark and Reed (1988), Capozza and Sick (1991) and Williams (1991) model conversion decisions where the density of development is the decision variable. Capozza and Li (1994) have constructed a more general model that accounts for the capital intensity of the development, property taxes and spatial variations in the pattern of property prices. Trigeorgis (1996, section 11.4, pp 349-356) reviews a specific class of models that examine the valuation of operating leases with options for early cancellation, extending the life of the lease and purchase of the lease.

Downing and Wallace (2000) consider an extension of the Dixit and Pyndyck model (1994) to account for stochastic volatility. A calibrated version of this extended model is used heuristically to identify influences over residential investment in the US. The authors assume that housing prices are a function of a vector of attributes, each generating a flow of services to the homeowner. Rental rates are set in a competitive market subject to demand shocks that evolve stochastically as geometric Brownian motions. The user cost of capital is the risk-free spot rate of interest adjusted for depreciation and costs of repair. In addition, the instantaneous risk-free rate evolves in accordance with a stochastic volatility model. Parameter values for this interest rate model are calibrated in accordance with a range of estimates from related empirical studies. Theoretical predictions from the calibrated real options model for an ascribed range of attribute values—changes in the spread between the rental rate and spot rate of interest, and in the volatility of the spread—are then compared with estimates taken from a mixed logit regression model of housing investment decisions to determine the validity of the theoretically inspired findings. The latter regression is estimated using panel data from the American Housing survey over the period 1985-1997 (Downing and Wallace, 2000; section 3)[3].

In a simple two period setting, Abel has extended the options pricing approach to investment to accommodate varying degrees of investment reversibility. He shows that the naïve net present value rule can only be applied after the cost of purchasing an additional unit of capital has been adjusted to take into account: the negative cost of extinguishing the marginal call option to purchase that same unit in the following period; and, the positive cost of acquiring a marginal put

[3] The variables in the regression model include volatility and spread variables, a user cost of capital that includes maintenance costs, federal and state marginal tax rates and property taxes, per capita income to

option to sell that additional unit in the following period (Abel, 1995). Abel goes on to demonstrate the relationship holding between the option pricing approach, the user cost of capital, and marginal q.

Nevertheless, over and above the issue of stochastic variations in the risk-free rate, there are additional limitations in the Dixit and Pyndyck framework. These include the fact that it accounts neither for the effects of transaction costs, nor for the effects of stochastic volatility in the underlying asset. These inadequacies are the major focus of this paper. However, the discussion inevitably leads to the consideration of notions of uncertainty that are broader than those examined into the existing literature on options pricing. For example, within the specialized literature on asset pricing the implications of Knightian or Keynesian uncertainty is currently a rapidly expanding area of active research models. One inroad into these issues is via applications of risk-sensitive and robust stochastic control theory, a topic that I shall examine in the next section of the paper.

3.0 Finance Applications of Risk-sensitive and Robust Control

In the asset pricing literature, techniques of risk-sensitive and robust control under stochastic uncertainty constraints have also been utilized to account for uncertainty, transaction costs, stochastic volatility in the underlying asset, and stochastic interest rates. I now intend to review aspects of this technically demanding literature. My intention is to provide an essentially heuristic and descriptive overview of risk-sensitive and robust control techniques so that the reader can gain an intuitive appreciation of what has motivated the various finance applications. In particular, I focus on certain appealing and useful characteristics of *risk-sensitive* objective functions.

Applications of risk-sensitive and robust control and filtering principles to finance theory are less common than macroeconomic applications to optimal stabilization policy. Nevertheless, notable exceptions include Lefebvre and Montulet's (1994) utilization of risk-sensitive, calculus-of-variations techniques to investigate a firm's optimal choice of the mix between liquid and illiquid assets, Fleming's (1993) risk-sensitive approach to portfolio management, and McEneaney's (1997) work on robust pricing of financial options under stochastic volatility. McEneaney's

---

account for business cycle effects, a variable for the average age of the house, and dummy variables to account for length of tenure and time-on-market.

robust control approach enables him to easily derive the Black and Scholes price for a standard Ito process. For a deterministic process, he derives a price corresponding to the conventional stop-loss hedging procedure, while for a stochastic volatility process, he demonstrates that a sufficient hedge is provided by the Black and Scholes price for the upper bound over the stochastic volatility process.

Risk-neutral control techniques only attend to the mean and variance of the relevant series rather than to higher-order moments and moments about the mean. It is easy to confirm that exponential objective functions are sensitive to all relevant moments within the joint-probability distribution. For example, Caravani begins with the conventional Linear Quadratic Gaussian (LQG) objective function:

1. $$\min \sum_{t=0}^{T} \{x'(t)Qx(t) + u'(t)Ru(t)\} \quad , \qquad \text{s.t. } x(t+1) = Ax(t) + Bu(t) + w(t)$$

Here *x* and *w* are *n*-vectors, *u* is an *m*-vector, *A* and *B* are conformable matrices and *t* is discrete time. The matrices *A* and *B* are assumed to be known through estimation, while *w* is assumed to represent random shocks over the relevant control period 0≤*t*£*T*. Furthermore, he assumes that:

1. Q is positive semidefinite, R is positive definite

2. $x(0)$ is known

3. $Ew(t) = 0$, $Ew(t)\, w(s) = Wd\,(t-s)$,

He shows that for the LQG case, a unique solution exists with the feedback form $u(t) = -G(t)x(t)$, where $G(t)$ is the solution to the relevant Ricatti equation. Caravani (1987, p 456) then examines two risk-sensitive H_2 Norms[4] e.g.:

2. $$f(x) = x + m|x| \quad , \qquad f(x) = \frac{1}{2}x[1 + \exp(mx)]$$

He demonstrates that in the scalar case:

[4] The H_2 norm is defined in the accompanying appendix.

3. $Ef'(x)Qf(x) = g(x_0) + \frac{\partial g}{\partial x}\Big|_{x_0} E(x - x_0) + \frac{1}{2!}\frac{\partial^2 g}{\partial x^2}\Big|_{x_0} E(x - x_0)^2 \ldots,$

It can be seen that the Taylor's expansion represents an infinite sequence of moments of increasing order[5].

Hansen, Sargent and Tallarini (1999) apply risk-sensitive control to a Lucas-style equilibrium asset-pricing model, whose dividend stream is derived from an entirely separate habit-persistence model of consumption. Uncertainty about the stochastic process is captured by a norm bound over the external perturbation term in the state variable equation. Adopting a "multiple-priors" interpretation, Hansen, Sargent and Tallarini argue that the non-uniqueness implied by the stochastic constraint depicts a form of Knightian uncertainty. Ambiguity of beliefs is not fully specified in probabilistic terms but is instead described by a set of specification errors, with a range defined by the imposed norm bound. Drawing on the equivalence between risk sensitive control and the Kalman filter *under certain limiting conditions*, the authors are able to apply standard, Kalman-filter based estimation techniques for the linear quadratic Gaussian case[6]. From this estimate they determine the magnitude of the respective uncertainty premia that might be embodied in asset prices, by considering variations in the magnitude of the risk-sensitivity parameter (and in the magnitude of another key parameter that reflects the size of the exogenous

---

[5] In Bielecki and Pliska's portfolio-theoretic application their chosen objective function is (Bielecki and Pliska, 1999, p 339):

$$J_q = \liminf_{t\to\infty} \left(\frac{-2}{q}\right) t^{-1} \ln Ee^{-(q/2)\ln V(t)}, \quad q > -2, q \neq 0 \cdot$$

By taking a second-order Taylor's series expansion about $q = 0$ the authors confirm that:

> […] J_q can be interpreted as the long run expected growth rate minus a penalty term, with an error that is proportional to q^2. Furthermore, the penalty term is proportional to the *asymptotic variance* … (p. 339).

Hence, maximizing J_q protects an investor interesting in maximizing the expected growth of their capital against large deviations of the actually realized rate from their expectations, where q plays the role of the risk-aversion parameter and $R(t) = \ln V(t)$ is the cumulative reward. This relationship reappears, in a different garb, in Stutzer's duality results discussed in section 5.0.

[6] Boel, James and Petersen (1997, p.11) establish that their risk-sensitive estimator reduces to the Kalman filter when the underlying model is linear with Gaussian error terms, exponential-of-quadratic cost and when the cost does not include accumulated error, thus, only penalizing the current time. Any departures from these assumptions imply that the Kalman filter lacks robustness in regard to model uncertainty, observation error and external perturbation. They consider a simple, scalar linear system with uncertainty in the drift of the state equation showing that the risk-sensitive filter has good performance (measured in

shocks to preferences in their habit-persistence based model of consumption). Andersen, Hansen and Sargent (1999) generalize these results to the continuous-time case and also allow for more general forms of uncertainty (e.g. perturbations to the transition probabilities of a hidden Markov model). To this end, they draw upon techniques of risk-sensitive control under relative entropy constraints[7].

In his review of the macroeconomic planning framework developed by Hughes-Hallet and Rees (1983) Brandsma (1986) observes that an extension of the optimal stabilization problem to one predicated on *minimisation of the variance* of the objective function introduces the second and third moments into the expression for the optimal control. He focuses on two characteristics of the resulting outcome. The first involves the application of more (less) weight or relative penalty on high (low) risk variables that exhibit large (small) variance. The second aspect involves a further scaling of each target vector in proportion to the relative priority each element possessed in the original quadratic objective function. Brandsma comments on the fact that:

> […] mean-variance decisions that optimise the expectation of a second-order approximation to a von Neumann-Morgenstern utility function fit into the expected utility analysis of Machina (1982). But the difficulty remains that the underlying utility function and probability distributions are unknown. From a practical point of view it is therefore attractive to be able to treat measures of uncertainty as an amendment to the preferences in their original quadratic form—which was itself an approximation to a more general objective—without having recourse to make too many extra assumptions (Brandsma, 1986, p 304).

---

terms of the mean square error) under nominal (i.e. zero error) conditions and acceptable performance that degrades much less rapidly than the Kalman filter

[7] The latter set of constraints—an alternative form of stochastic uncertainty constraint to norm bounds over model error, observation error and external perturbation—can be interpreted in terms of the discrepancy between a reference probability distribution and the unknown actual probability distribution. Andersen, Hansen and Sargent's paper (1999) call this constraint the entropy bound The duality holding between relative entropy and free entropy allows the authors to define error bounds over the minimum risk-sensitive estimator (see Boel, James and Petersen, 1997, for a more complete exposition of this material).

This modified quadratic approach is also one adopted by Aaron Tornell (2000) and applied to an asset-pricing problem. Tornell adopts an H^{∞} control framework using a conventional quadratic utility function to interpret well known asset-pricing anomalies. While the H_2-norm utilized in linear quadratic gaussian (LQG) control applications minimises the *rms* values of the regulated variables when the disturbances are unit intensity white-noise processes, the H^{∞}-norm is bounded from below by the *rms* gain of the system[8]. Thus, the H^{∞}-norm minimises the worst-case *rms* value of the regulated variables when the disturbances have unknown spectra. This deterministic perturbation of unknown spectral dimension represents the combined effects of the observation error, external disturbance, and any intrinsic (multiplicative) model uncertainty. The robustness parameter(g^2), then reflects the controller's sensitivity towards these various sources of uncertainty.

Pointedly, Tornell rejects the view that asset-pricing anomalies are a result of misperception or irrationality (p. 1). While continuing to accept the assumption that, knowing the model that generates payoffs, agents must filter the persistent and transitory components of a sequence of observation in order to estimate the unobservable state of the economy, his point of departure is that agents are not perfectly sure about this model. For one thing, the disturbance process may be misspecified and, for another thing, the payoff model, itself, may be improperly formulated. In their risk-sensitive approach Hansen, Sargent and Tallarini's (1999), assume that shocks are normally distributed. In contrast, uncertainty is not parameterized in Tornell's analysis. His H^{∞} approach models uncertainty in the form of unknown disturbance sequences with a bounded l_2-norm. Thus, while the rational expectations solution to the control problem is designed to achieve the best performance conditional on the absence of misspecification in the relevant probability distribution, the H^{∞} solution to the control problem is designed to perform well under any norm-bounded misspecification. Moreover, while under rational expectations forecasts (that are based on a recursive version of the Kalman filter) can be formed independently from agent's choices, under H^{∞} control, forecasts and robust portfolio choices are *jointly* determined. Critically, Tornell no longer assumes that the state is perfectly observed. Observation error is both an intrinsic and also an important part of the story he wants to tell about asset pricing. In the limit, when the robustness parameter g approaches infinity, the H^{∞} forecasting formulas coincide with their

---

[8] In practical terms, the H^{∞}-norm is based on the singular value decomposition of the state system's transfer function matrix into maximal and minimal eigenvalues (see the accompanying appendix and Shahian and Hassul, 1993, p. 446).

rational expectations counterparts. In particular, the recursive estimator for the state is replaced by the conventional Kalman filter[9].

Tornell shows that regressions of excess returns on observed dividend price ratios reveal no significant predictability when the sequences derived from rational expectations formulas are employed, whereas estimates from the H^{∞} formulas, with low values for the robustness parameter g, exhibit significant predictability in conformity with the empirical evidence (Table 1, p. 30). In addition, Tornell demonstratess that, for low values of the robustness parameter, the H^{∞} prices also tend to violate the variance bounds, whereas the rational expectation prices do not. These variance bounds are calculated by comparing the variance of price estimates with the variance of price sequences that would have prevailed if the discount factor was constant and agents had perfect knowledge of future dividends (Table 2, p. 32). Tornell calculates 100 dividend sequences and tabulates the number of times the variance of the forecast price sequences exceeds that of the perfect foresight price sequences. In addition, Tornell presents data on the magnitude of the equity premium and the risk-free rate for various values of the g parameter. As expected, the equity premium increases as g falls. For $g = 0.5$ the premium attains the empirically plausible value of 5.9% while maintaining an equally plausible low risk-free rate of 1.49% (Table 3, p. 34). Finally, Tornell examines how well the rational expectations and H^{∞} formulas track actual US stockprices over the period 1871-1996. The H^{∞} price sequences track the actual S&P500 index much better than their rational expectations counterparts. Tornell attributes the ability of H^{∞} prices to exhibit similar anomalies to the actual US data to the fact that H^{∞} forecasts are more sensitive to dividend news (and as we have seen, this sensitivity is inversely related to the size of the robustness parameter). He concludes:

---

[9] In contrast to LQG control, Tornell shows that the H^{∞} *certainty equivalence principle* breaks the original problem into three sub-problems. First, for a given value of the state variable, portfolio and consumption strategies are derived through *backward* dynamic programming. In this initial stage a sequence of disturbances is selected from the set that is compatible with observed dividends that is the *worst possible* given the objective function for the dynamic game. Second, the agent solves a *forward* dynamic programming problem to extract the persistent component from past dividend's observations conditional on the state variable. At this stage it transpires that the sequence of unknown disturbances in the state equation has no bearing on the chosen optimum. However, the optimal amounts of each asset are chosen to maximize the objective function, given that nature has already selected the worst possible sequence of observation errors. Third, H^{∞} estimates of the state variable, the observation variable, and equilibrium asset prices are derived using the value functions associated with forward and backward dynamic programming solutions to the previous two sub-problems.

> The point we want to make is that, in a simple asset pricing model, excess sensitivity to news can results from either: (a) *misperception* of the duration of shocks in a behavioural setup, or (b) from a desire for *robustness* in an H_∞ setup. (Tornell, 2000, p. 38)

A key difference is that in an H_∞ approach, agents effectively use the same 'nominal' model as RE agents, whereas in a behavioural setup agents employ a different model. Tornell reviews some of the approaches taken in this behavioural literature (p. 38-9), observing that in some models agents are *overconfident* about the precision of their private signals (i.e. they perceive the noise-to-signal ratio to be lower than it actually is), or else, the noise-to-signal ratio is reduced through the *association* of current events with memories evoked of similar past events, while in other models agents overreact to several signals pointing in the same direction because they project trends whereas, in fact, the earnings series is presumed to follow a random walk[10].

In defence of the robust or risk sensitive control approach I would point out that an enormous amount of intellectual power and technical virtuosity is applied to analyzing outcomes in financial markets so that well-resourced investors can profit from the misperceptions of others[11]. The robust approach to asset pricing has the advantage that it does not have to rely on what is effectively only a set of simplistic metaphors for representing various forms of agent misperception. In contrast, I would contend that the concern for robustness is firmly grounded in *substantive* and *persistent* ontological features of the control and filtering environment[12]. In addition, within a utility maximizing framework, risk-sensitive value functions can be derived

---

[10] We could add to this review the literature on *Rational Belief Equilibria* (Kurz and Motolese, 1999), in which agents estimate regime switching models with fewer dimensions than exist in actuality. Nevertheless, the unconditional moments of the estimates conform closely to those of the real world. In this case, a particular form of bounded rationality is responsible for agent misperception.

[11] Of course, the noise-trading literature responds to this argument by contending that more rational agents must operate within short-term investment horizons that prevent them from taking positions in assets that are known to be mis-priced but will take too long to return to their fundamental values. In contrast, in the literature on Adaptive Rational Expectations Dynamics (Brock and Hommes, 1997), investors switch between high performance and high cost rational estimators and low-cost low performance myopic estimators based on a calculation of the relevant trade-off between performance and cost which can vary depending on how closely or far away the system is from the steady-state.

[12] The ubiquitous influence of coordination failure across markets, technological change, non-constant returns to scale, and changes in income distribution between classes and between various fractions of capital, springs to mind.

from plausible and empirically supported extensions to the axioms of either the von-Neumann Morgenstern or the Savage classes of expected utility theory (see Epstein and Zin, 1989).

4.0 Option Pricing and General Equilibrium Asset Pricing under Complete Markets

An excellent resource on options pricing is Cochrane's (2000) book on asset pricing. Cochrane begins with a simple derivation of the fundamental asset pricing equation relating the appropriately discounted expected payoff x_{t+1} to the asset's price p_t (p. 15):

4. $$p_t = E_t\left[b\,\frac{u'(c_{t+1})}{u'(c_t)}x_{t+1}\right] = E_t\left[m_{t+1}x_{t+1}\right]$$

The key variable in his analysis is the stochastic discount rate m_{t+1}. In accordance with convention, Cochrane demonstrates that either the law of one price or non-arbitrage is sufficient to prove the existence of such a discount rate (see Cochrane, 2000, Chapter 4). In findings that are tied together in Chapter 6, Cochrane demonstrates that both the minimum variance return situated on the conventional mean-variance frontier and also the beta representation of the CAPM can each be formally linked to, and derived from, the asset pricing equation (4), appearing directly above.

When it comes to the pricing of options the traditional approach used in deriving the Black-Scholes (1972) formula is one based on the construction of a portfolio of stocks and bonds that replicates the instantaneous payoff of the option. Under the law of one price, the price of the option and the price of the replicating portfolio must be equal. After some algebraic manipulation using Itô's lemma, this equivalence gives rise to a partial differential equation that can be solved for the option price. An alternative approach, increasingly in vogue and one that Cochrane prefers, draws upon the stochastic discount factor formulation.

At each date the option is priced using the discount factor *m*, that prices both the stock and the bond. For a call option, Cochrane (2000, sections 17.2.1-17.2.2) shows that one can either solve the discount factor forward and then find the call option value by using $C = E(mx^C)$, or characterize the price path for the option using Itô's lemma and solve it backwards from expiration. This sort of analysis is now standard fare in up-to-date finance texts and will not, therefore, be examined in further detail. However, I shall review the pricing of options in incomplete markets—a field that is very much the focus of current research—in more detail.

5.0 Option Pricing and General Equilibirium Asset Pricing under Incomplete Markets

During extreme events such as stockmarket crashes, when continuous trading is impossible (events that are frequently modeled as Poisson jump processes in finance theory) or when interest rates and stock volatility are stochastic, then the law of one price breaks down. In these conditions a replicating portfolio of securities—one that provides a perfect hedge against the corresponding shocks—cannot be constructed (see Cochrane, 2000, Chapter 18).

This is one area of finance where techniques of robust control have been applied with some measure of success. As discussed in the previous section, McEneaney (1996) has used a robust control approach to price options where volatility is stochastic and bounded. It transpires that the option price is one generated by the familiar Black-Scholes formula with a constant volatility equal to the upper bound over the volatility[13]. His paper is technically demanding, employing viscosity solutions for the resulting Isaacs equations, and less mathematical readers would probably appreciate a more heuristic or intuitive description of what is involved in pricing options when markets are incomplete.

Cochrane sets out his own "good deal bounds" solution to the problem of imperfect replication in Chapter 18 of his text following Cochrane and Saá-Requejo (2000). The good deal bounds are derived by "systematically searching over all possible assignments of the "market price of risk" of the residual, constraining the total market price of risk to a reasonable value, and imposing no arbitrage opportunities, to find upper and lower bounds on the option price" (p. 301). The residual

---

[13] McEneaney (p.15) assumes that the asset price process is a continuous semi-martingale driven by a Brownian motion process as given by the stochastic differential equation:

$$dP_t = bP_t dt + s\,(P_t Y_t)P_t dB_t$$

while the wealth of the contingent claims writer is given by:

$$dX_t = [rX_t + (b-r)l_t]dt + s\,(P_t, Y_t)\,l_t dB_t$$

where the volatility of the asset price process is dependent on another, independent, stochastic process:

$$dY_t = f_1(Y_t) + f_2(Y_t)dB_t^{(2)}.$$

Here, $B^{(2)}$ is another Brownian motion on the given filtration (**W**,**F**,***P***) independent of B, r is the riskfree rate of return, $\boldsymbol{l}$ is the amount of wealth invested in the underlying asset at any given time, $\sigma(p,y)\in [0, \sigma_2]$ where $\sigma_2 < \infty$ is the relevant upper bound, f_1 is Lipschitz and $f_2 \in C^4{}_b$. The differentiability assumptions on f_2 can be relaxed.

to which Cochrane refers is the error term in a projection of the option payoff onto the space of portfolio payoffs that can be constructed from the basis assets (a stock and a bond in the conventional Black and Scholes framework).

Essentially, the option is priced using a dynamic recursive programming approach maximizing the expected discounted payoff to the option under a series of one-period constraints (for example, for a European call the objective function becomes: $\max_{\{m\}} E(mx^c),\ x^c = \max(S_T - K, 0)\cdot$ Here, S_T is the price of the underlying stock while K is the strike price. One constraint imposes non-arbitrage (ie. the stochastic discount rate must be non-negative), another characterizes the asset pricing model that has been used to price the underlying stocks and bonds (p = E[mx]), while the last imposes a bound over the Sharpe ratio of the imperfectly replicating portfolio[14]. By restricting the range of discount factors to those falling within a sphere around the origin determined by a volatility constraint that takes the form $E(m^2) \leq A^2$, a range of option prices is generated that are narrower than the conventional arbitrage bounds. The positivity constraint obviously rules out negative prices that are not ruled out merely through the imposition of the volatility bound. Cochrane demonstrates how to solve the multiple-period problem using Kuhn-Tucker techniques to arrive at the relevant upper and lower bounds over the option price.

However, following the work of Bernardo and Ledoit (2000) Cochrane observes (p. 318) that an alternative approach could be taken that, instead, would entail the imposition of bounds over the ratio of the gain ($[R^e]^+ = \max(R^e, 0)$) and the loss ($[R^e]^- = -\min(-R^e, 0)$) of the excess return[15]. Cochrane emphasizes the exact analogy holding between the key duality relationship that

[14] The Sharpe ratio is the absolute value of the excess return of an asset divided by the return variance. Cochrane's approach draws on the duality between discount factor volatility and the Sharpe ratios established by Hansen and Jagganathan (1991). Using the covariance decomposition, the following identity can be derived from the asset pricing relationship for excess returns R^e:

$$0 = E(mR^e) = E(m)E(R^e) + \boldsymbol{r}_{m,R^e}\boldsymbol{s}(m)\boldsymbol{s}(R^e).$$

Because $|\boldsymbol{r}| \leq 1$ and $E(m) = 1/R^f$ when a riskfree rate R^f exists, the following relation must hold between the Sharpe ratio and discount factor volatility:

$$\frac{\boldsymbol{s}(m)}{E(m)} \geq \frac{|E(R^e)|}{\boldsymbol{s}(R^e)}.$$

From a geometrical analysis of this relationship Hansen and Jagganathan derive the precise duality relationship between discount factor volatility and Sharpe ratios that Cochrane compares to Bernardo and Ledoit's gain-loss ratio, as discussed immediately below.

[15] It should be noted that the asymmetry between up-side and downside risk is a notable feature of Bielecki and Pliska's (1999) continuous-time, risk-sensitive portfolio model.

Bernardo and Ledoit derive, which compares the gain-loss ratio to the ratio of the supremum and infinum of the stochastic discount rates[16]:

5. $$\max_{\{R^e \in \mathbf{R^e}\}} \frac{E[R^e]^+}{E[R^e]^-} = \min_{\{m: 0=E(mR^e)\}} \frac{\sup(m)}{\inf(m)},$$

and the well known Hansen and Jaganathan duality relationship between discount factor volatility and the Sharpe ratio:

6. $$\max_{\{R^e \in \mathbf{R^e}\}} \frac{|E(R^e)|}{s(R^e)} = \min_{\{m: 0=E(mR^e)\}} \frac{s(m)}{E(m)}.$$

He notes that this analogy "…hints at an interesting restatement of asset pricing theory in L^1 with sup norm rather than L^2 with second moment norm." (p. 318).

The gain-loss ratio summarizes the attractiveness of a zero-cost investment for the benchmark investor. When it equals unity the investment is fairly priced for the investor, but if it exceeds unity the benchmark investor would receive more gain than necessary for him or her to increase holdings in that asset. If markets were complete, the set of pricing kernels $E(mR^e)$ that correctly price all portfolio payoffs would have a unique element; otherwise it would have many elements. Using an asterisk to designate equilibrium outcomes using the benchmark pricing kernel, Bernardo and Ledoit's duality result can be framed in terms of deviations from the benchmarking pricing kernel (Bernardo and Ledoit, 2000, p. 151):

$$\max_{\{R^e \in \mathbf{R^e}\}} \frac{E*[R^e]^+}{E*[R^e]^-} = \min_{\{m: 0=E(mR^e)\}} \frac{\sup\limits_{j=1,\dots,S}\left(\dfrac{m_j}{m_j*}\right)}{\inf\limits_{j=1,\dots,S}\left(\dfrac{m_j}{m*_j}\right)},$$

where $j = 1,\ldots,S$ is the relevant state space and the expectation $E*$ is taken under risk-adjusted probabilities:

$$p_j* = p_j u'(c_j*)/E[u'(\mathbf{c}*)]; \quad j = 1,\ldots,S \;.$$

[16] In the draft of Cochrane's text that I have cited, the expectations operators appear to have been erroneously left off in the ratio appearing on the left hand side of the depicted equation.

Thus the gain is the expectation of the excess payoff computed over those states in which the excess payoff is positive[17]. Bernardo and Ledoit see the main advantage of their gain-loss duality result being the fact that it characterizes the set of arbitrage and approximate arbitrage opportunities. The existence of a bound on the gain-loss ration is equivalent to imposing the restriction that a ³ m/m^* ³ b so as to sharpen the non-arbitrage restriction that ¥ ³ $m/m^* > 0$ (i.e. the maximum gain-loss ratio $\overline{L}$ is just the ratio β/α). If the benchmark model is reasonable, then high gain-loss investments are inconsistent with well-functioning capital markets. If $\overline{L}$ increases (decreases) this reflects less (more) confidence in the ability of the benchmark model to price non-basis assets. In the limit, as $\overline{L}$ approaches infinity, the bounds over the price of the chosen non-basis asset will approach the non-arbitrage bounds, and as $\overline{L}$ approaches 1, the bounds for an option would approach the Black-Scholes price.

Bernardo and Ledoit (p. 168) observe that several other duality results in the literature could, in principle, be used to derive asset price bounds. For example, the Hansen and Jagannathan bounds can be generalized to account for restrictions on the kth moment of the pricing kernel. However, Bernardo and Ledoit also cite research by Stutzer (1995), who demonstrates that a restriction over the maximum expected utility attainable by an investor whose preferences conform to *constant absolute risk aversion* is equivalent to a restriction on the entropy of the pricing kernel $E[m\log(m)]$. Stutzer's duality result enables Bernardo and Ledoit to link their gain-loss approach to the related literature on entropy measures and Bayesian estimation. As we have seen, entropy also features in risk-sensitive control theory, where it appears as in the form of a relative entropy constraint.

---

[17] In a continuous–time setting Bernardo and Ledoit show that it can be calculated by decomposing the expression for the gain into three or fewer terms, each representing a linear function of the stock price over an interval (if the strike price K falls within an interval that interval must be broken into two subintervals at K). Each term can be computed given the interval bounds S_1 and S_2, using the formula:

$$E^*\left[(a + bS)\mathbf{1}_{[S_1 \le S \le S_2]}\right] = a\left[\Phi\left(d_1 - s\sqrt{t}\right) - \Phi\left(d_2 - s\sqrt{t}\right)\right] + bSe^{rt}\left[\Phi(d_1) - \Phi(d_2)\right],$$

where $d_i = \dfrac{\log\left(S/S_i e^{-rt}\right)}{s\sqrt{t}} + \dfrac{1}{2}s\sqrt{t}, \quad i = 1,2$ and a and b are the coefficients of the linear function. This gives the gain (and by symmetry, the loss) for any portfolio weights w_S and w_C, given initial prices S and C. The authors next show that by setting each of the weights to:

$$w_S = wE^*\left[\mathbf{C} - e^{rt}C\right];\ w_C = 1 - wE^*\left[\mathbf{S} - e^{rt}S\right], \text{ where } \mathbf{C} = (\mathbf{S} - K)^+$$

the free parameter can now be varied and the portfolio value $E^*[R^e]$ will stay constant. The value of w can then be chosen to minimize the first absolute moment (L_1 norm) of the excess payoff—a straightforward univariate convex optimization problem. The final stage requires the imposition of bounds on the maximum gain-loss ratio. Bounds on the option price can be derived by inverting the maximum gain-loss function with respect to its argument in C.

6.0 Stutzer's Entropy-Based Analysis of Asset-Pricing Models

Stutzer (1995) introduces the unconditional moment conditions that are defined by an asset pricing model over gross real returns R^i for each of the I assets, under the state probability measure m (p. 369):

7. $E[R^i m] = \int R^i m dm = 1, \quad i = 1,\ldots, N$

He then introduces the risk-free asset with unit real payoff $X_t^0 = 1$, gross return (i.e. gross real interest rate) r, and, therefore, time-varying price $1/r$. In the usual way, he next obtains the unconditional moment condition for this asset, namely:

8. $E[m] = E[1/r] = c$, where c is the fixed mean.

He then proceeds to derive the familiar Hansen and Jagganathan (1991) affine benchmark—an affine combination of excess returns that delivers the minimum variance amongst those stochastic discount factors (SDF) satisfying the above two moment conditions (Stutzer, 1995, p. 371):

$$m^a = (R - E[R])' w^a + c \,.$$

Where w^a is a vector of coefficients. On substitution into the first moment condition, the resulting expression can be solved for the unique vector w^a (when invertibility conditions are satisfied). This vector also determines the particular affine combination, m^a amongst all possible combinations $m(w)$ given by the preceding expression, that is closest to any SDF m satisfying the moment conditions, in the sense of mean squared distance. In addition, it is the vector of weights determining the mean-variance efficient portfolio.

Stutzer arrives at another version of the moment conditions by dividing the first by the second introducing a measure change ($dv = m/E[m]dm$) that enables him to establish the following equivalence between expectation operators (p. 374):

$$9. \quad E\left[R^i \frac{m}{E[m]}\right] = \int R^i \frac{m}{E[m]} dm = \frac{1}{c} = \int R^i dv = E_v\left[R^i\right].$$

Stutzer derives a variational characterization of the set of *risk neutral measures*, v^l, which satisfy the latter expectational relation and have a *state price probability density* (SPD), dv/dm. He achieves this result by minimizing the relative entropy or *Kullback-Leibler Information Criterion* (KLIC) $I(v,m)$ given by (p. 375):

$$10. \quad v^l = \arg\min_v \ I(v,m) \equiv \int \log(dv/dm) dv,$$

over the set of of SPDs satisfying equation (9) above. Under appropriate regularity conditions, Stutzer demonstrates that, associated with this convex problem is the Gibbs density[18]:

$$11. \quad \frac{dv^l}{dm} = \frac{\exp\left(\sum_{i=1}^N w_i^l R^i\right)}{E\left[\exp\left(\sum_{i=1}^N w_i^l R^i\right)\right]}$$

Stutzer provides four interpretations of the Gibbs SPD benchmark vector. First, he shows that the change of measure relative to the candidate SDF m^c, with mean c from equation (9) also satisfies the following information bound inequality or minimum distance criterion (p. 376):

$$12. \quad I(v^l, m) \le E\left[\frac{m^c}{E[m^c]} \log\left(\frac{m^c}{E[m^c]}\right)\right] = I(v^c, m)$$

This inequality plays the same role as Hansen and Jagganathan's variance bound inequality. Second, he provides a quasi-maximum likelihood interpretation of the benchmark (pp. 377-378). Stutzer's quasi-maximum likelihood interpretation of the affine benchmark portfolio and associated Gibbs density is based on the *directed orthogonality property* (Stutzer, p. 377)

---

[18] The Gibbs SPD is derived from the first order conditions for the following problem:

$$w^l = \arg\min_{w_1, w_2, \dots w_N} M(w) \equiv E\left[\exp\left[\sum_{i=1}^N w_i^l \left(R^i - 1/c\right)\right]\right],$$

dividing through by $\exp\left(-\sum_{i=1}^N w_i^l c\right) E\left[\exp\left(\sum_{i=1}^N w_i^l R^i\right)\right]$,

and solving for 1/c, confirming the fact that the Gibbs SPD satisfies the requisite moment conditions. The information bound can, itself, be calculated from the same problem as $I(v^l, m) = -M(w^l)$.

satisifed by the KLIC criterion[19]. Stutzer's version of the principle of directed orthogonality is given below:

$$v^l \equiv \arg\min_V \, I(v, m) \left| \int R_i \frac{m}{E[m]} dm = \frac{1}{c} = \arg\min_{V(W)} \, I(v, v(w)). \right.$$

A series of straightforward manipulations of $I(v, v(w))$ yields the following:

$$I(v, v(w)) = \int \ln \frac{dv}{dv(w)} dv = \int \ln \frac{dv/dm}{dv(w)/dm} \equiv E_m \left[\frac{dv}{dm} \ln \frac{dv}{dm} \right] - E_v \left[\ln \frac{dv(w)}{dm} \right].$$

Because the first term in the above identity is independent of w, selecting a vector v that maximizes the second term minimizes $I(v, v(w))$. Thus, the Gibbs benchmark SPD v^l (based on the benchmark portfolio weights w^l) that minimizes the constrained KLIC criterion $I(v, m)$ (i.e. subject to the relevant moments conditions) also minimizes the unconstrained KLIC criterion $I(v, v(w))$[20].

---

[19] The directed orthogonality principle is similar to the orthogonality of best approximation in Hilbert space for generalized inverse solutions to underdetermined systems of linear equations, where the latter linear equations are the constraints, the distance measure employed is the squared Euclidean and the prior estimate is the zero vector. Jones (1989) shows that directed orthogonality is a natural extension of this problem to the task is one of specifying a positive measurable function on a domain of positive measure subject to a series of linearly independent, locally bounded integral constraints. In this more general case, the prior P is an estimate of the density arrived at without knowledge of the linear functional values (constraints) and for admissible functions Q, the solution is chosen by minimizing the directed distance between members in the class Q and the prior P.

[20] Csiszár (1985) clarifies the point that Akaike's Bayesian approach differs markedly from applications of the extended maximum entropy principle to the derivation of risk-neutral measures (note that in Cziszár's notation Q stands for the prior distribution and $D(\cdot \mid \cdot)$ represents the distance or maximum entropy function):

> Akaike considers statistical estimation problems, adopting the view that an unknown distribution (rather than a parameter) is to be estimated, and he uses $D(P \mid Q)$ as a loss function measuring the loss when the unknown true distribution P is estimated by Q. The maximum entropy principle, as understood in this paper, relates to problems of a different kind, not within the scope of standard statistical decision theory, namely to updating priors to conform to evidence typically consisting in moment constraints. Although in both cases $D(P \mid Q)$ appears as a measure of "distance" which should be minimized, a formal difference is that in Akaike's model minimization is performed with respect to the second variable while the maximum entropy principle calls for minimization with respect to the first one. (Csiszár, 1985, p. 98)

Third, Stutzer shows that the Gibbs benchmark weights w^l determine the composition of optimal portfolio for an investor with constant absolute risk aversion with a constant of proportionality equal to $-1/a$ where a is the coefficient of absolute risk aversion[21].

Fourth, Stutzer offers a Bayesian interpretation of the Gibbs benchmark SPD based on axiomatic arguments that the KLIC-based density minimizes the information gained by a change of measure satisfying the moment conditions, without incorporating any extraneous information. In other words, when m is uniformly distributed (a common representation for the Bayesian prior distribution under uncertainty) the solution v^l of problem (17) is a Bayesian posterior update of the prior risk-neutral distribution in the light of sample information consistent with the relevant moment conditions (16).

Shore and Johnson (1980), in work which is cited in Stutzer's paper, provide an alternative axiomatic justification for both maximum entropy and minimum cross-entropy based on a set of four reasonable principles of statistical inference: thus departing from earlier justifications that are instead based on the information-theoretic properties of entropy measures[22]. Without going too far into the technical detail of their analysis, these axioms include: *uniqueness* (i.e. for a specified prior and for new information restricted to a set that includes at least one density with

---

[21] For a CARA investor with initial wealth W_0, who invests $\sum_{i=1}^{N} w_i$ in risky assets and the remainder in the riskless asset, terminal wealth would be:

$$W = W_0\left(\frac{1}{c}\right) + \sum_{i=1}^{N} w_i \left[R_i - \frac{1}{c}\right]$$

The associated optimal utility can therefore be determined from:

$$V \equiv \max_{w_1, w_2, \ldots, w_N} E[U(W)] = -e^{-\frac{aW_0}{c}} E\left[e^{\sum_{i=1}^{N} a w_i \left(R_i - \frac{1}{c}\right)}\right] = U\left(W_0 \frac{1}{c}\right) E\left[e^{\sum_{i=1}^{N} -a w_i \left(R_i - \frac{1}{c}\right)}\right].$$

Now $U(W) < 0$ but marginal utility is positive. Therefore, the maximum of V is equivalent to minimization of the second term. When compared with the Gibbs benchmark SPD it can be seen that $w^* = -w^l/a$. Stutzer also derives a utility-based characterization of the information bound:

$$\left|\log\left(\frac{V}{U(W_0/c)}\right)\right| = I\left(v^l, m\right).$$

[22] Henri Theil's (1974) approach to *Rational Random Behaviour* exploits the information-based interpretation of Jayne's entropy measure. Here, entropy functions as a measure of likelihood, which allows Bayesian decision-makers to arrive at a posterior distribution by multiplying their prior distribution by the entropy measure. Using a calculus of variations approach, Theil shows that the appropriate measure is constructed by taking the exponent of the (negative of the) ratio of the loss function (i.e. the loss arising from a decision based on the incorrect control variable) over the marginal cost of information (given by Jaynes entropy). For a lucid overview see Theil, 1978, pp. 255-261.

finite distance measure, the resulting posterior should be unique); *invariance* (a change in coordinate system that can be represented by a transformation with an invertible Jacobian should not matter to the result); *system independence* (it should not matter whether one accounts for independent information about independent densities separately in terms of different prior densities, obtaining separate posterior densities, or together in terms of a joint density for the prior, because there should be no interaction between the two systems); and *subset independence* (it should not matter whether one treats an independent subset of system states in terms of a separate conditional density or in terms of the full system density in obtaining the posterior density).

7.0 Kitamura and Stutzer's Entropy-based GMM Framework

Kitamura and Stutzer (1997) have developed an alternative to the optimal minimum distance (OMD) estimator first proposed by Hansen and Singleton (1982) for generalized method of moments (GMM) estimation. The former estimator, based on minimisation of the Kullback-Leibler Information Criterion, is asymptotically as efficient as the latter OMD, has the same data requirements and computatonal feasibility, but is more efficient in small samples. The OMD is biased in small samples because sampling errors in the second moment are correlated with sampling errors in the estimate of the covariance matrix of the sample moments. Kitamura and Stutzer use their estimator to construct a χ^2-specification test of the moment conditions as well as Wald, Legrange multiplier and Likelihood ration tests of parametric restrictions, analogous to those commonly used in applications of OMD.

Kitamura and Stutzer closely follow the presentation in Hansen (1982), commencing with a stochastic vector process $\boldsymbol{x}_t$, t = 1,2,…, a parameter vector b from a set Q of possible parameter vectors, and an r-component vector of observable, real-valued functions $f(\boldsymbol{x},b) = (f_1,\ldots,f_r)'$. The authors denote the observed time-series by $f(x_1,b),\ldots,f(x_T,b)$. Theory is represented by the prediction $E^m\left[f\left(\mathbf{x},b^*\right)\right] \equiv \int f\left(\mathbf{x},b^*\right)dm\left(\mathbf{x}\right) = \mathbf{0}$, where b is a parameter vector from Q, E^m is the expectation with respect to probability measure m, and $\mathbf{0}$ denotes an r-component vector of zeroes. Empirical content is given to the theoretical representation by assuming that: $\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} f\left(x_t,b^*\right) \equiv \mathbf{0}$, for most realizations of the process. Hansen's GMM estimator of

b^* satisfying theoretical priors is achieved by finding $\hat{\boldsymbol{b}}$ which makes the observed vector of sample means $\hat{f}_t(b) \equiv (1/T)\sum_{t=1}^{T} f(x_t, b)$ close to **0**. Specifically:

13. $\hat{b} = \arg\min_{b \in \Theta} \tilde{f}_T(b)' W \tilde{f}_T(b)$.

As Hamilton shows (1994,Chapter 14), the weighting matrix W is calculated from the inverse of the asymptotic variance $\boldsymbol{S}$, where the latter given by:

14. $\mathbf{S}_T^* \equiv (1/T)\sum_{t=1}^{T} [f(x_t, b_0)][f(x_t, b_0)]' \xrightarrow{P} \mathbf{S}$.

An iterative procedure is required using an arbitrary weighting matrix such as $\mathbf{W}_T = \mathbf{I}_r$, to estimate b_0, which is then used in the above expression to produce a new estimate of $\mathbf{W}_T = [\mathbf{S}^{(0)}{}_T]^{-1}$, for use in deriving a new estimate of b. The iterative procedure is repeated until a convergence criterion is met. The Newey-West adjusted-estimate of **S** can be used if there is serial correlation in the $[f(x_t, \boldsymbol{b})]_{t=-\infty}^{\infty}$ process.

Hamilton (1994) demonstrates that the GMM technique is sufficiently general to embrace a range of other econometric models as special cases, including: ordinary least squares, instrumental variable estimation, estimators for systems of non-linear simultaneous equations, and dynamic rational expectations models. For example, OLS estimation implies the following set of orthogonality conditions:

15. $E[\mathbf{x}_t(y_t - \mathbf{x}_t'\boldsymbol{b}_0)] = 0$.

The terms inside the brackets conform to what is required of the set of $f(\boldsymbol{x},b)$ functions. For instrumental variables estimation the relevant r-vector of orthogonality conditions becomes: $E[\mathbf{x}_t(y_t - \mathbf{z}_t'\boldsymbol{b}_0)] = 0$, where $\mathbf{z}_t$ is a vector of explanatory variables, and $\mathbf{x}_t$ is a vector of predetermined explanatory variables that are correlated with **z** but uncorrelated with u_t, the residual vector from the regression model $y_t = \mathbf{z}_t'\boldsymbol{b} + u_t$. For non-linear systems of simultaneous equations of the form:

$\mathbf{y}_t = \mathbf{g}(\mathbf{x}_t, b) + \mathbf{u}_t$, where $\mathbf{z}_t$ is a $(k \times 1)$ vector of explanatory variables, and b is an $(a \times 1)$ vector of unknown parameters, the orthogonality conditions can be expressed in the required form:

16. $$\mathbf{f}(\mathbf{x}_t, b) = \begin{bmatrix} [y_{1t} - f_1(\mathbf{x}_t)]\mathbf{z}_{1t} \\ [y_{2t} - f_2(\mathbf{x}_t)]\mathbf{z}_{2t} \\ \vdots \\ [y_{nt} - f_n(\mathbf{x}_t)]\mathbf{z}_{nt} \end{bmatrix}.$$

Here $\mathbf{z}_{it}$ is a vector of instruments that are uncorrelated with the ith element of $\mathbf{u}_t$. Finally, for estimating dynamic rational expectations models, the relevant set of Euler equations becomes the set of real-valued functions that is to be estimated through GMM techniques (see Hamilton, 1994, pp. 416-24 for details).

Kitamura and Stutzer's proposed replacement for the Hansen estimator is one based on a non-linear projection problem:

17. $$\min_{P \in \mathbf{P}(b)} D(P : m) \equiv \min_{P} \int \log(dP/dm)\,dP$$
$$\text{subject to } E^P[f(\mathbf{x}, b)] = \mathbf{0}$$

where $D(P : m) \equiv \int \log(dP/dm)\,dP$ is the Kullback-Leibler Information Criterion distance from P to m. The optimal estimator b is found by making $D(P{:}m)$ as close to zero as possible. Kitamura and Stutzer (1997, pp. 864-5) show that this is achieved by finding the saddlepoint of the following function:

18. $$\mathbf{M}(b, g) \equiv E^m\left[e^{g f(\mathbf{x}, b)}\right],$$

where $g(b) = \arg\min_{g} \mathbf{M}(b, g)$, and $b^* = \arg\max_{b} \mathbf{M}(b, g(b))$. As we have seen above, this function arises naturally from the expression for the *Gibbs canonical density* (p. 864):

19. $$\frac{dP(b)}{dm} = \frac{\exp\left[g(b)' f(\mathbf{x}, b)\right]}{E^m\left[\mathbf{exp}\left[g(b)' f(\mathbf{x}, b)\right]\right]}.$$

Under a series of standard assumptions (p. 866), the authors show that an asymptotically efficient estimator can be calculated by replacing the observation $f(\mathbf{x}_t, b)$ with:

20. $$\hat{f}(t, \boldsymbol{b}) \equiv \sum_{k=-K}^{K} \frac{1}{2K+1} f(x_{t-k}, \boldsymbol{b}),$$

where $K^2/T \to 0$ and $K \to \infty$ as $T \to \infty$[23]. The estimator is then determined by:

21. $$\left(\hat{\boldsymbol{b}}_T, \hat{g}_t\right) = \arg\max_{b} \min_{g} \left[\hat{Q}_T(\boldsymbol{b}, g) = \frac{1}{T} \sum_{t=1}^{T} e^{g'} \hat{f}(t, \boldsymbol{b})\right].$$

In a further discussion of other practical asset pricing problems that can be addressed using their approach Bernardo and Ledoit note that:

> Real options are difficult to value using arbitrage methods since the stochastic component of the options return often cannot be replicated because the underlying asset does not exist, does not trade, trades in an illiquid market, or is not spanned by a portfolio of traded assets. If one can construct an imperfect hedging strategy by using some combination of existing assets, then our gain-loss restriction yields bounds consistent with the inability to construct extremely attractive portfolios using these basis assets. (p. 168)

This simple observation underlies what I am attempting to achieve in this paper. I have argued that the gain-loss and entropy approaches to options pricing mirror the findings of researchers who have used techniques of risk-sensitive and robust control to price assets. However, for deriving effective bounds on option prices in incomplete markets, the former body of literature affords more straightforward methods—that could, for example, be incorporated into spread-sheet or lattice-based models—with results are, in essence, similar to those that can only be attained

---

[23] These assumptions primarily relate to characteristics of the $\mathbf{x}_t$ process that guarantee asymptotic normality (i.e. strong mixing, stationary and ergodic), but they also include moment existence conditions, differentiability of $f(x, \boldsymbol{b})$ at the optimum, non-singularity of the denominator in the Gibbs canonical density, and continuity and uniqueness of the $\boldsymbol{b}$ parameter satisfying the constraints (1).

with much greater effort and more demanding levels of technical virtuosity through the application of via robust control techniques. This characteristic of entropy-based methods will be demonstrated in the next section of the paper.

8.0 Minimum Cross-Entropy and Martingale Measures: the Discrete-Range Case

The relationship between minimum entropy measures and martingales can be seen most clearly in the discrete-time and discrete-range case. Unfortunately, a different notation is required for the discrete range case because expectations and constraints have to be defined over a finite partition of the probability space. Consider a filtration $\mathbf{F_N} = \mathbf{F} = \sigma(P_N) = \{\omega_1, \omega_2, \ldots, \omega_N\}$ representing the evolution of the information process as a random sequence $\{A_t\}$ of subsets of the finite sample space Ω. Moreover, the P_t are sequences of partitions of Ω such that for each block A in P_t there exist blocks $A_1, \ldots, A_k$ in P_{t+1} such that $A = \cup A_j$. Also, for any $\omega \in \Omega$ there exists a sequence of blocks $B_N\{\omega\} = \{\omega\} \subset B_{N-1}\{\omega\} \subset \ldots . B_0\{\omega\} = \Omega$, each contained in P_{t+1}. Any probability Q on (Ω, F) is determined by a sequence of conditional probabilities $Q[\cdot \mid A_t]$ for A_t in P_t and $t = 0,1,\ldots,N$ as follows:

$$Q\{w\} = Q_{B_0\{w\}}(B_1\{w\})Q_{B_1\{w\}}(B_2\{w\})\cdots Q_{B_{N-1}\{w\}}(B_N\{w\}).$$

When defined over such a filtration, we know that a risk-neutral probability measure Q turns S^*_t into an (F_t, Q) martingale that satisfies the following equation:

$$22.\quad E_Q\left[S_n^*(t+s)\middle|F_t\right] = E_Q\left[J_t\, S_n(t+s)/J_{t+1}\middle|F_t\right] = S_n^*(t);\quad t,s \geq 0.$$

Here $J = \{J_t : t = 0,1,\ldots,T\}$ is the bank account process with $J_0 = 1$ and the risk-free interest rate $r_t = (J_t - J_{t-1} / J_t) \geq 0$; while $S_n = \{S_n(t): t = 0,1,\ldots,T\}$ is the non-negative stochastic process representing the time t price of the nth risky security and $S^*_n(t) \equiv S_n(t)/ J_t$, $t = 0,1,\ldots,T;$ $n = 1,2,\ldots,N$ is the discounted security price. Since the S^*_t vector is F_{t-1}–measurable its components are constant on the blocks of P_{t+1}. Let $S^*_{t+1} = \mathbf{y_j}$ on A_j. Assume that each B_j is partitioned into blocks $B_{kj} \in P_t$, for $j = 1,2,\ldots,M(j)$; $M(j) \geq 1$ and $\Sigma M(j) = N(t)$. Keeping t fixed, let the values of S^*_t on the block B_{ij} be denoted by $\mathbf{y}_{ij} \in \Re d$ with $\mathbf{s}_j$ denoting the value of S^*_{t-1} on the block B_j. In this context, Henryk Gzyl (2000) develops a risk neutral measure through the application of the maximum entropy principle. First, he introduces the usual discrete-range non-arbitrage conditions:

23. $\sum_{j=1}^{k} \mathbf{y}_j q_j = \mathbf{s}, \quad \sum_{j=1}^{k} q_j = 1,$ with the F_{t+1} measurable prices $S^*_{t+1} = \mathbf{y_j}$ on the jth partition A_j .

The convex class P(p,**s**), assumed to be non-empty, is then defined in relation to these conditions as follows:

24. $P(p,\mathbf{s}) = \{q(j) = a(j)p(j) | a(j) > 0;\ j = 1,\dots,k; \sum q(j) = 1;\ \sum \mathbf{y_j} q(j) = \mathbf{s}\}$

Second, a concave function $S_p(q)$ is defined over this convex space as follows:

$$S_p(q) = -\sum q(j) \ln a(j).$$

This function is arrived at through the substitution of $q(j)$ and $p(l,j)$ for the functions $f(j)$ and $g(j)$ in the discrete-range version of the Kullback-Leibler formula $K_p(f,g) = \sum f(j) \ln(f(j)/g(j)) p(j)$. Third, Gzyl defines the following exponential family (obviously related to the Gibbs SPD), parameterized by $l \in \Re^d$:

25. $p(l,j) = \dfrac{e^{-\langle l, y(j)\rangle} p(j)}{E_p\left[e^{-\langle l, \mathbf{Y}\rangle}\right]} = \dfrac{1}{Z(l)} e^{-\langle l, y(j)\rangle} p(j),$

where the $\mathbf{y_j}$ are values of the random variable **Y**, such that P(**Y = y**$_\mathbf{j}$) = p(j). It must be he case that:

$S_p(q) \le \Sigma(l) := \ln Z(l) + \langle l, \mathbf{s}\rangle$. As $Z(l)$ is convex on $\Re^d$, and as $\Sigma(l) \to \infty$ when $\|l\| \to \infty$, then $\Sigma(l)$ achieves its unique minimum at a certain point l^* in $\Re^d$. Moreover, at this point it can be confirmed that $p(l^*) \equiv p^*$ is in P(p,**s**), and at this point of minimum entropy $S_p(p^*) = \Sigma(l)$ (Gyzl, 2000, p. 6; Csiszár, 1985, pp. 86-87)[24].

[24] Incidentally, Imre Csiszár (1985) calls the updated probability density derived by minimizing the relative entropy or directed divergence (as reflected in the Kullback-Leibler number for all admissible densities) the "*I*-projection" and provides a Bayesian justification for its use that is similar to that offered by Stutzer.

Pulling together the all the previous strands of analysis, it must be the case that the requisite risk neutral measure Q on (Ω, F) can now be constructed from its sequence of conditional expectations with respect to the filtration $\{F_t;\ t = 1,\ldots,N\}$ in accordance with:

26. $$Q_{ij} = Q_{B_j}(B_{ij}) = \frac{e^{-\langle l_t(j), y_{ij}\rangle}\mathbf{1}}{Z(t,\ l_t(j))}\, p_{ij}$$

Because $Z(t,\ l_t(j)) = \sum_{B_{ij}\subset B_j} e^{-\langle l_t(j), \mathbf{y}_{ij}\rangle} p_{ij}$, then $p_{ij} = P(B_{ij})/P(B_j)$ can be replaced with $P(B_{ij})$. Moreover, $l_t(j)$ will be constant on each B_j on F_{t-1} as will $Z(t,\ l_t(j))$, so that these variables are both F_{t-1}-measurable. Thus:

27. $$r_t := \frac{e^{-\langle l_t, S_t^* - S_{t-1}^*\rangle}}{\sum_k e^{-\langle l_r(j), (\mathbf{y}_{ij} - \mathbf{s}_j)\rangle} p_{kj}} = \frac{e^{-\langle l_t, \Delta S_t^*\rangle}}{Z_\Delta(t,\ l_t)\big|_{B_j}} = \frac{e^{-\langle l_t, S_t^*\rangle}}{Z(t,\ l_t)},$$

since S^*_{t-1} is constant on the blocks B_j of P_{t-1}. Moreover:

28. $$\Sigma(t,\ l_t) = \ln Z(t,\ l_t) + \langle l_t, S_{t-1}^*\rangle = \ln Z_\Delta(t,\ l_t).$$

Now consider an F_{t-1}-measurable function H. The risk-neutral expectation of H can be calculated in accordance with the following:

29. $$E_Q[H|F_{t-1}]I_{B(j)} = E_Q[H|B_j] = \Sigma_k H(k) q_{kj} = \sum_{B_{ij}\subset B_j} H(k)\frac{e^{-\langle l_t(k), y_{kj}\rangle}}{Z(t,\ l_t,(j))}P(B_{kj}) = \int_{B_j} r_t H dP \cdot$$

Thus,

$$E_Q[H|F_{t-1}] = \frac{E_p[rH|F_{t-1}]}{E_p[r|F_{t-1}]} = E_p\left[\frac{\Psi_t}{\Psi_{t-1}} H \middle| F_{t-1}\right],\ \text{where } \Psi_t = E_p[r|F_{t-1}] \text{ and } r = \prod_{t=1}^{N}(\Psi_t/\Psi_{t-1}) = \Psi_N/\Psi_0 \cdot$$

Here, $\Psi_0=1$, so that $r = \Psi_N$.

Accordingly, $r_t = \Psi_t / \Psi_{t-1} = \dfrac{e^{-\langle l_t, \Delta S_t^* \rangle}}{Z_\Delta(t, l_t)}$, and the function $\Sigma_\Delta(t,\lambda_t)$: $\Re^d \times \Omega \rightarrow \Re$ defined above has a minimum at $\lambda_t(\omega) \in \Re^d$, for each $\omega \in \Omega$, $\lambda_t : \Omega \rightarrow \Re^d$ is F_{t-1}-measurable, and (Gzyl, 2000, theorem 3.1, p 10):

30. $$r = \exp\left\{-\sum_{t=1}^{N}\left[\left\langle l_t, \Delta S_t^* \right\rangle + \ln Z_\Delta\left(t, l_t\right)\right]\right\},$$

defines a density for Q with respect to P such that S^*_t is an (F_t, Q)-martingale.

Gzyl (pp. 11-13) uses this result to derive risk-neutral measures for the jump probabilities in a binomial, trinomial and a generic Markov chain model[25]. At this juncture, two specific aspects of Shore and Johnson's research (1980), which I briefly review in the appendix, are pertinent. First, for the discrete-range case, they establish that maximum entropy is a special case of minimum cross-entropy when the linear constraints are known to be binding, but no prior is available to the researcher (Shore and Johnson, 1980, p. 33). Second, using the same axiomatic properties, they demonstrate that minimum cross-entropy is still the appropriate distance measure to adopt, even for cases where the linear constraints that are to be imposed take the form of *inequalities* rather than *equations* (Shore and Johnson, 1980, equation 4, p. 43). This latter result provides a direct link to efforts by researchers such as Cochrane (2000) and Bernardo and Ledoit (2000), who endeavour to extend option pricing to the case of incomplete markets by imposing range bounds over the Sharpe ratio or gain-loss ratio. It also confirms that Stutzer's minimum cross entropy approach to the diagnosis of asset-pricing models can, likewise, by generalized to the case of incomplete markets. Rather than imposing a bound over the gain-loss ratio of the form a ³ m/m^* ³ b , as do Bernardo and Ledoit, the researcher can impose a pair of inequality bounds directly over the non-arbitrage conditions[26].

---

[25] Also, see Smith and Nau (1995) for an example of real option pricing using the binomial lattice model with a CARA utility function.

[26] Kapur and Kesavan (1994, pp. 344-345) consider the original Markovitz portfolio choice problem, where each of the given expected returns on each security can be viewed as a moment condition, supplemented by non-negativity conditions on the amounts of each security in the portfolio and the requirement that portfolio shares must sum to unity. They demonstrate that if the Havrada-Charvat measure of minimum cross entropy is employed as an alternative to variance in this problem, then minimizing this measure is equivalent to maximizing the expected utility of a person with a CARA utility function of the form, $u(x) = x^{1-a}$. This result establishes a clear entropy-based link between Stutzer's stochastic discount rate approach to estimating the SDF and the earlier literature focusing on the Markovitzian minimum-variance frontier.

9.0 Conclusion

In this paper I have focused on entropy as the unifying vehicle for investigating aspects of real options theory under incomplete markets. Stutzer's (1995) research provided a range of alternative interpretations of the benchmark, state price probability density in asset pricing models. I then examined Kitamura and Stutzer's application of entropy-based techniques to GMM estimation of asset pricing models. Finally, I reviewed Gzyl's (2000) use of entropy-based methods to derive martingale measures, which could then be applied to binomial, trinomial and Markov-chain models to price both real and financial options. I drew on material that examined entropy-based estimation for cases where inequality constraints were imposed over the requisite measure changes. This provided a clear link between the entropy techniques that I discussed and the new research on "good-deal" bounds and norm-bounds over gain-loss ratios for the pricing of options in incomplete markets. In future research I intend to utilize these entropy-based techniques to value various kinds of real and financial option.

Later in the same chapter (p. 351), they observe that a particular case may arise when the researcher may wish to impose inequality constraints over the actual probabilities that a state may arise. If, for the discrete-range case, one wishes to minimize cross entropy subject to the constraints:

$$\sum_{i=1}^{n} p_i = 1; \quad \sum_{i=1}^{n} p_i g_{ri} = a_r, \quad r = 1,2,\ldots,m; \quad a_i \le p_i \le b_i,$$

then Kapur and Kesavan recommend the adoption of the following generalized measure of minimum cross entropy:

$$\sum_{i=1}^{n} (p_i - a_i) \ln \frac{p_i - a_i}{q_i - a_i} + \sum_{i=1}^{n} (b_i - p_i) \ln \frac{b_i - p_i}{b_i - q_i}.$$

This measure is more convenient than the KLIC since the probabilities obtained automatically satisfy the inequality constraints unlike the former. To apply this technique to Stutzer's problem let $q_i = m_i$ and $p_i = v_i$. The moment constraints can then be written as:

$$\sum_{i=1}^{n} p_i g_{ri} = a_r \Rightarrow \sum_{i=1}^{m} R_i \frac{m}{E(m)} m_i = \frac{1}{c} = E_v[R_i]$$

The inequality constraints thus impose bounds over the discrete-range version of the Radon-Nikodym derivative $dv = m/E[m]dm$.

Boel, James and Petersen (1997) use relative entropy to calculate error bounds for the more general case in which the actual state is observed with error. They assume that the true probability model is fixed, but unknown, and that the estimation procedure makes use of a fixed nominal model. They demonstrate that the resulting error bound for the risk-sensitive filter is the sum of two terms, one of which coincides with an upper bound on the error one would obtain if one knew exactly the underlying probability model, while the other is a measure of the distance between the true and design probability models. In the particular option pricing case under investigation here, the divergence represented by each of these two terms has been replaced by another pair of terms: one representing the measure change required to convert the stochastic return process into a martingale and the other representing the inequality bound over the resulting Radon-Nikodym derivative (i.e. $m/E(m)$, the stochastic discount rate). In both cases this divergence is measured by the relative entropy metric, but in the former case relative entropy is implicitly embodied in the exponential characterisation of the risk-sensitive filter, while in the latter case it is embodied in the set of inequality constraints over the relevant moments that are reflected in a set of induced bounds over the Gibbs benchmark state price probability density (i.e. via the KLIC or generalized entropy measure). If Kapur and Kesavan's (1994) generalized entropy measure is employed as a substitute for the KLIC, then ordinary Lagrangian multiplier techniques can be used to solve the minimum cross entropy problem rather than the more demanding Kuhn-Tucker techniques because it possesses the advantage of automatically satisfying the inequality constraints.

In the option pricing literature it is typically assumed that the individual investor can exert no influence over the stochastic properties of the underlying asset. In such cases it is difficult to see why a change in the degree of uncertainty aversion would influence the dimensions of inequality constraints obtaining over the generalized minimum cross-entropy measure. However, if the option valuation process were embedded in a more comprehensive model, with a monetary asset made available as a potential hedge against uncertainty over prospective returns, then one could readily associate rising uncertainty aversion with a rising preference for liquidity as real investments are postponed or abandoned[27].

Further research is progressing in the field of options pricing in incomplete markets that will have direct relevance to real options theory. For example, Howe and Rustem (1997) have developed improved numerical algorithms for solving infinite range minimax optimization problems that can be applied both to individual option pricing and to portfolio hedging strategies when investors face significant transaction costs. McEneaney (1997) foreshadows similar research that he intends to conduct utilizing techniques of robust control. A property developer can readily be conceived of as managing a portfolio of both real and financial options, and the transactions costs associated with hedging real options exposure are often sizeable. Accordingly, this field of research into minimax hedging would potentially have practical benefits.

With the increasing availability of high frequency data, evidence is accumulating that financial time-series exhibit the fractal characteristics that are associated with non-linear chaotic dynamics (Dacorogna et al, 1993). Once again, robust control techniques have been employed for filtering non-linear systems. Einecke and White (1999) have developed a robust version of the Extended Kalman Filter (EKF). In their model norm bounds governing model uncertainty represent linearization errors in the second-order EKF. Techniques of this nature have obvious applications in finance for cases where the underlying dynamic system is known to be non-linear.

Tornell (2000) envisages further developments in finance-related applications of robust control that allow for time-variation in the robustness parameter over the business cycle. These variations would reflect changing investor sentiment or uncertainty aversion as the economic environment improves or deteriorates. Treating uncertainty aversion as an endogenous parameter seems to violate the neoclassical penchant to treat changes in preferences and technology as exogenous to the system. However, it mirrors related research into *adaptive belief systems* (Brock and Hommes, 1997), which presumes that investors are able to switch predictors in response to endogenous changes in their relative cost and predictive performance.

---

[27] Klaus Nehring (1999) has developed an axiomatic basis for choice that exhibits a "preference for flexibility" under uncertainty, in the sense that the agent wants to keep her options open so that she can respond to *anticipated* but *unforseen* contingencies. In future, this sort of framework could potentially be applied to portfolio choice and real options theory, in incomplete markets, where the money asset operates as a hedge against such forms of uncertainty.

The inverse relationship between the robustness parameter g and the risk-sensitive parameter q, a parameter which also appears within the related expression for the entropy integral, confirms the affect that uncertainty aversion has on outcomes irrespective of whether the relevant decisions were modeled using stochastic risk-sensitive control, maximum entropy or deterministic robust control techniques. I have established that such endogenous fluctuations in the robustness parameter are directly related to the entropy-based alternative to the variance bounds, good-deal bounds, or the gain-loss ratio. A change in uncertainty aversion alone, quite apart from any variation in the stochastic characteristics of the underlying assets, would result in a widening or narrowing of these bounds and ratios.

Further progress in this direction affords the exciting prospect of combining neoclassical finance theory with Keynesian research into uncertainty, liquidity preference and animal spirits as influences over investment behaviour. As uncertainty increases, and prices move further along the continuum away from a unique reference value more towards the non-arbitrage bounds over the value of the real options, the spread between the "bid" and "ask" prices that reflect "market incompleteness" would widen. As a result, investment activity would inevitably decline. This is the ultimate insight that I have been striving for in this paper.

Bibliography

Abel, A. B., Avinash K. Dixit, J. C. Eberly and Robert, S. Pyndyck (1996) "Options, the Value of Capital, and Investment", *Quarterly Journal of Economics,* Vol. 104, No. 3, pp. 753-77.

Avellaneda, Marco, Freidman, Craig, Holmes, Richard and Samperi, Dominick (1997) "Calibrating volatility surfaces via relative entropy minimization," *Applied Mathematical Finance*, Vol. 4, pp. 37-64.

Bernardo, A. and Ledoit, O. (1999) "Gain, Loss and Asset Pricing", *Journal of Political Economy*, Vol. 108, February, pp. 144-172.

Bielecki, T.R.; Pliska, S.R. (1999) "Risk-sensitive dynamic asset management," *Applied Mathematics and Optimization*, Vol.39, No. 3, May - June, pp.337-360.

Black, F. and Scholes, M. S. (1973) "The pricing of options and corporate liabilities," *Journal of Political Economy*, ol. 81, May/June, pp. 637-54.

Boel, R. K., James M. R. and Petersen, I. R. (1997) "Robustness and Risk-Sensitive Filtering" Research Mimeo funded by ARC and Belgian Program on InterUniversity Attraction Poles.

Bossaerts, P., Ghysels, E., and Gouriéroux, C. (1996) "Arbitrage-based pricing when volatility is stochastic," DP #9641, CREST.

Brock, W. A. and Cars H. Hommes (1997) "A Rational Route to Randomness," *Econometrica*, Vol. 65, No. 5, September, pp. 1059-1095.

Capozza, D. R. and Helsey R. (1990) "The stochastic city" *Journal of Urban Economics*, Vol. 28, No. 2, September, pp 295-306.

Capozza, D. R. and Sick, G. A. (1991) "Valuing long-term leases: the option to redevelop," *Journal of Real Estate Finance and Economics*, Vol. 4, No. 2, June, pp. 209-223.
Clarke, H. and Reed, W. (1988) "A stochastic analysis of land development, timing and property evaluation," *Regional Science and Urban Economics*.
Cochrane, John H. and Saá-Requiro, J. (2000) "Beyond arbitrage: good-deal asset price bounds in incomplete markets," *Journal of Political Economy*, Vol. 108, February, pp. 70-119.
Cochrane, John H. (2000) *Asset Pricing*, http://www-sb.uchicago.edu/fac/john.cochrane/research/papers, book draft, June 12.
Cox, D. (1981) "Statistical analysis of time series: some recent developments" *Scandinavian Journal of Statistics*, Vol. 8, pp. 93-115.
Csiszár, I. (1991) "Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems", *The Annals of Statistics*, Vol. 19, No. 4, pp. 2032-2066.
Csiszár, Imre (1985) "An Extended Maximum Entropy Principle and a Bayesian Justification," in Bernardo, J. M.., De Groot, M. H., Lindley, D. V., and Smith, A. F. M., (Eds) *Bayesian Statistics 2*, pp. 83-98, Elsevier Science Publishers, B. V., North Holland.
Dacorogna M. M., Müller, U. A., Nagler R. J., Olsen R. B., and Pictet O. V., (1993) "A geographical model for the daily and weekly seasonal volatility in the FX market," *Journal of International Money and Finance*, **12** (4), pp. 413-438.
Dixit, Avinash K. and Robert S. Pyndyck (1994) *Investment and Uncertainty*, Princeton, Princeton University Press).
Einicke, G. A. and White, L. B. (1999) "Robust extended Kalman filtering" *IEEE Transactions on Signal Processing*, Vol. 47, No. 9, September, pp. 2596-2599.
Engle, R. and Russell, J. (1998) "Autoregressive conditional duration: a new model for irregularly-spaced transaction data" *Econometrica*, Vol. 66, pp. 1127-1162.
Fleming, W. H. (1995) Optimal investment models and risk-sensitive stochastic control, in Davis, M., et al., ed. *Mathematical Finance*, Springer-Verlag, New York, pp. 75-88.
Föllmer, H. and Schweizer, M. (1991) "Hedging of contingent claims under incomplete information," *Applied Stochastic Analysis, Statistics Monographs*, Volume 5, eds. M. H. A. Davis and R. J. Elliott, Gordon and Breach, London/New York, pp. 389-414.
Glover, K. and Doyle, J. C. (1988) "State-space formulae for all stabilizing controllers that satisfy an H_∞-norm bound and relations to risk sensitivity," *Systems and Control Letters,* Vol. 11, pp. 167-172.
Green, M. and D. J. N. Limebeer (1995) *Linear Robust Control*, Prentice Hall, New Jersey
Gzyl, Henryk (2000) "Maxentropic construction of risk neutral measures: discrete market models" Memo, May, Venezuelan School of Mathematics, Caracas, Venezeula.
Guiso, Luigi and Giuseppe Parigi (1996) "Investment and Demand Uncertainty" Banca D'Italia, *Temi di discussione del Servizio Studi*, Number 289, November.
Hansen, L. P. and Jagannathan, R. (1991) "Implications of security market data for models of dynamic economies", *Journal of Political Economy*, Vol. 99, April, pp. 225-262.
Howe, M. A., and Rustem, B. (1997) "A robust hedging algorithm," *Journal of Economic Dynamics and Control*, Vol. 21, pp. 1065-1092.
Hurn, A. S. and Robert E. Wright (1994) "Geology or Economics? Testing models of irreversible investment using North Sea Oil data" *The Economic Journal*, 104, March, pp. 363-371.
Iglesias, P. A., Mustafa, D., and Glover, K. (1990) "Discrete time H_∞ controllers satisfying a minimum entropy criterion," *System and Control Letters*, Vol. 14, pp. 275-286.
Kapur, J. N. and Kesavan, H. K. (1994) *Entropy Optimization Principles with Applications*, Academic Press, Boston.
Kitamura, Yuichi and Michael Stutzer (1997) "An information-theoretic alternative to generalized method of moments estimation" *Econometrica*, Vol. 65, No. 4, July, pp. 862-74.
Kurz, Mordecai and Maurio Motolese (1999) "Endogenous Uncertainty and Market Volatility" Working paper No. 99-005, University of Stanford, Economics Department, Working Paper Series, http//www-econ.stanford.edu/faculty/workp/index.html.
Leahy, John V. and Toni M. Whited (1996) "The effect of uncertainty on investment: some stylized facts" *Journal of Money, Credit and Banking*, Vol. 28, No. 1, February.
Lefebvre, M., and P. Montulet (1994) "Risk sensitive optimal investment policy," *International Journal of Systems Science*, Vol. 22, pp. 182-192.

Lei, David, Michael A. Hitt and Joel D. Goldhar (1996) "Advanced Manufacturing Technology: Organisational Design and Strategic Flexibility", *Organisation Studies*, 17/3, pp. 501-523.
Leland, H (1985) "Option pricing and replication with transaction costs," *The Journal of Finance*, Vol. 5, pp. 1283-1301.
McEneaney, W.M. (1997) "A robust control framework for option pricing" *Math. of Operations Research*, Vol 22, pp. 202--221.
Musiela, M. and Rutkowski, M. (1997) *Martingale Methods in Financial Modelling*, Springer Verlag, Berlin/Heidelberg.
Nehring, Klaus (1999) "Preference for flexibility in a Savage framework" *Econometrica*, Vol. 67, No. 1, (January), pp. 101-119.
Prigent, J. –L., Renault, O., and Scaillet, O. (1999a) "An autoregressive conditional binomial option pricing model," DP Université Catholoique de Louvain.
Prigent, J. –L., Renault, O., and Scaillet, O. (1999b) "Option pricing with discrete rabalancing," DP Université Catholoique de Louvain.
Russell, J., and Engle, R. (1998) "Econometric analysis of discrete-valued irregularly-spaced financial transaction data using a new autoregressive conditional multinomial model," CRSP Working Paper #470.
Schweizer, M. (1991) "Option hedging for semimartingales," *Stochastic Processes and their Applications*, Vol. 37, pp. 339-363.
Shahian, Bahram and Michael Hassul (1993) *Control System Design Using MATLAB â*, Prentice Hall, New Jersey.
Shore, John, E. & Johnson, Rodney, W. (1980) "Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy" *IEEE transactions on Information Theory*, Vol. IT-26, No. 1, January, pp. 26-37.
Smith, J. E. and Nau, R. F. (1995) "Valuing risky projects: option pricing theory and decision analysis," *Management Science*, Vol. 41, No. 5, (May), pp. 795-816.
Stutzer, M. (1995) "A Bayesian approach to diagnosis of asset pricing models" *Journal of Econometric Theory*, Vol. 68, August, pp. 367-97.
Theil, Henri (1978) *Introduction to Econometrics*, Prentice-Hall, New Jersey.
Theil, Henri (1974) "A theory of rational random behaviour," *Journal of the American Statistical Association*, Vol. 69, pp. 310-314.
Titman, S. (1985) "Urban land prices under uncertainty," A*merican Economic Review*, Vol. 75, No. 3, June, pp. 505-514.
Tornell, A (2000) "Robust-H_∞ forecasting and asset price anomalies" National Bureau of Economic Research Working Paper No. 7753, June, http://www.nber.org/papers/w7753
Trigeorgis, Lenos (1996) *Real Options: Managerial Flexibility and Strategy in Resource Allocation*, Cambridge, Massachusetts, The MIT Press, Chapter 3, "Option Pricing Theory and Financial-Options Applications"
Williams, J. (1991) "Real estate development as an option," *Journal of Real Estate Finance and Economics*, Vol .4, No. 2, June, pp. 191-208.

Technical Appendix:

Spectral Representations and Norms

This first section of the appendix introduces the 2-norm and H_∞-norms that are routinely applied in a control theory setting. For any series of numbers $\{x_t\}$ the Fourier transform is defined by:

$$x(w) = \sum_{t=-\infty}^{\infty} e^{-iwt} x_t \ .$$

This operation transforms a series that is a function of time into a complex-valued function of *w*. Given $x(w)$, we can recover x_t by the inverse Fourier transform:

$$x_t = \frac{1}{2\boldsymbol{p}} \int_{-p}^{p} e^{+iwt} x(\boldsymbol{w}) d\boldsymbol{w} = \frac{1}{\boldsymbol{p}} \int_{0}^{p} \left| x(\boldsymbol{w}) \right| \cos(\boldsymbol{w}t + \boldsymbol{f}(\boldsymbol{w})) d\boldsymbol{w}.$$

The above expression follows from the identity:

$$e^{iq} = \cos\boldsymbol{q} + i\sin\boldsymbol{q} \Rightarrow \cos\boldsymbol{q} = \left(e^{iq} + e^{-iq}\right)/2 .$$

The Euclidean or 2-norm of a matrix is defined by (Shahian and Hassul, 1995, pp. 442-3):

$$\left\|\mathbf{A}\right\|_2 = \max_{x \neq 0} \frac{\left\|\mathbf{A}x\right\|_2}{\left\|x\right\|_2} = \max_{\left\|x\right\|_2 = 1} \left\|\mathbf{A}x\right\|_2 = \bar{\boldsymbol{s}}(\mathbf{A}),$$

where the barred sigma notation stands for the largest singular value of the **A** matrix in the associated singular value decomposition (SVD). The SVD decomposes a rectangular matrix **A** with rank r into the product:

$$\mathbf{A} = \mathbf{U}\hat{\Sigma}\mathbf{V}^*, \quad \text{where}\, \mathbf{U}^*\mathbf{U} = \mathbf{I}_m, \;\; \mathbf{V}^*\mathbf{V} = \mathbf{I}_n,$$

$$\text{and}\; \hat{\Sigma} \begin{cases} [\Sigma \mid \mathbf{0}] \text{if } n > m \\ \begin{bmatrix} \Sigma \\ \hline \mathbf{0} \end{bmatrix} \text{if } n < m \end{cases} \quad where\, \Sigma = \begin{bmatrix} s_1 & & & \\ & s_2 & & \\ & & \ddots & \\ & & & s_r \end{bmatrix}, \; r = \min\{m, n\}$$

Here, the eigenvalues are subscripted in ascending order of magnitude. Therefore, the largest singular value $\bar{\boldsymbol{s}}$ is then defined as the highest eigenvalue, s_1.

For signals or time functions $x(t)$, the *2*-Norm the *2*-norm defined by:

$$\left\|x(t)\right\|_2^2 = \int_0^\infty x(t)^2 dt = \frac{1}{2\boldsymbol{p}} \int_{-\infty}^{\infty} \left|X(j\boldsymbol{w})\right|^2 d\boldsymbol{w},$$

can be interpreted as the gain of the system. When the above norm is finite, the function is said to belong to the L_2 Hilbert space (ie. it is square integrable). If a matrix is conceived as a system with x as its input and Ax as its output, then the 2-norm represents the maximum gain.

The H_2- Norm for the *single-variable* transfer function G is defined by

$$\|G\|_2^2 = \frac{1}{2p}\int_{-\infty}^{\infty} \left|G(jw)\right|^2 dw.$$

The power spectral density of the system output is given by:

$$\mathrm{S_y}(w) = \left|G(jw)\right|^2 S_x(w).$$

Hence, the root mean squared value of the output is given by:

$$y_{rms} = \left\{\frac{1}{2p}\int_{-\infty}^{\infty}\left|G(jw)^2 S_x(w)dw\right|\right\}^{\frac{1}{2}}$$

Shahian and Hassul observe that for white noise inputs, $S_x(w) = 1$ for all frequencies (p.446). Therefore, the H_2-Norm can be interpreted as the RMS value of the output when the system is driven by white noise input. Similarly, they argue that for a *single-variable system* the H_∞-norm defined by:

$$\|G\|_\infty = \sup_{x \neq 0}\frac{\|Gx\|_2}{\|x\|_2} = \sup_{w}\left|G(jw)\right|,$$

must satisfy the following inequality bound:

$$\|G\|_\infty \geq \frac{\left\{\int_{-\infty}^{\infty} S_y(w)dw\right\}^{\frac{1}{2}}}{\left\{\int_{-\infty}^{\infty} S_x(w)dw\right\}^{\frac{1}{2}}} = \frac{y_{rms}}{x_{rms}}.$$

In other words, the H_∞-Norm is bounded from below by the *rms gain* of the system. In the *multivariable* case the H_2- and H_∞-Norms are defined by:

1. $\|G\|_2^2 = \frac{1}{2p}\int_{-\infty}^{\infty} trG*(jw)G(jw)dw = \frac{1}{2p}\int_{-\infty}^{\infty}\sum_i s_i^2[G(jw)]dw$

and,

2. $\|G\|_\infty = \sup_w \bar{s}[G(jw)]$, respectively.

In summary, the H_∞-norm minimizes the worst-case *rms* value of the regulated variables when the disturbances have unknown spectra, whereas the 2-norm minimizes the *rms* values of the regulated variables when the disturbances are unit intensity white-noise processes.

LQG, H-infinity Control, and Maximum Entropy

The following section of the appendix summarises the somewhat remarkable link between stochastic LQG control problems, deterministic H_∞-problems and maximum entropy. Glover and Doyle (1988, p. 170) consider a discrete-time, risk sensitive, LQG stochastic control problem for the state equation:

3. $x_t = Ax_{t-1} + B_1 w_t + B_2 u_{t-1}$

4. $z_t = C_1 x_{t-1} + D_{11} w_t + D_{12} u_{t-1}$

5. $y_t = C_2 x_{t-1} + D_{21} u_{t-1}$

where the process/observation noise *wt* is white and Gaussian with unit variance. The relevant quadratic cost function is:

6. $G = x_T^* \Pi x_T + \sum_{t=0}^{T-1} z_t^* z_t$.

The risk sensitive optimal controller minimizes:

7. $g_T(q) = -\frac{2}{qT}\log E[\exp(-qG/2)]$.

For a stabilizing linear time invariant (LTI) controller with transfer function $K(s)$ connected from y to u, and cost function $g_T(q)$, the output z_t will be a stationary Gaussian process with spectrum:

$$f\,(\,\boldsymbol{l}\,) = (1/2\boldsymbol{p})\mathbf{H}\left(e^{\,jl}\right)\mathbf{H}\left(e^{\,jl}\right)^{*}.$$

H is the closed loop transfer function $F_l(P,K)$ from w to z. Accordingly, $F_l(P,K) = P_{11} + P_{12}K(I - P_{22}K)^{-1}P_{21}$ and:

$$P(s) = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} \stackrel{s}{=} \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & \mathbf{0} & D_{22} \end{array}\right] = \begin{pmatrix} D_{11} & D_{12} \\ \mathbf{0} & D_{22} \end{pmatrix} + \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} (sI - A)^{-1} \begin{bmatrix} B_1 & B_2 \end{bmatrix}$$

Glover and Doyle introduce two lemmas which allow them to establish that (see lemmas 3.1, 3.2, p. 171):

8. $$\lim_{T\to\infty} g_T(q) = \begin{cases} \dfrac{1}{2pq}\displaystyle\int_{-p}^{p} \log\left(\det\left[\mathbf{I} + q\mathbf{H}\mathbf{H}^{*}\right]\right)dl & \text{if } q\left\|\mathbf{H}\right\|_{\infty}^{2} > -1 \\ \infty & \text{if } q\left\|\mathbf{H}\right\|_{\infty}^{2} < -1 \end{cases}.$$

In this expression the absence of the cost term $x_T^{*}\Pi x_T$ can be ignored since K is stabilizing and $T \to\infty$. Thus, any LTI optimal controller must be such that:

9. $$\left\|F_l(P,K)\right\|_{\infty} \le (-\boldsymbol{q})^{-1/2}, \qquad \text{for } \boldsymbol{q} < 0.$$

The integral in the preceding limit expression is the frequency domain version of the entropy integral, which can be minimized over all LTI controllers meeting the H_∞-norm bound. As q is made more negative the H_∞ control robustness parameter becomes larger. Beyond a certain critical point, all controllers would give infinite cost.

Howe and Rustem’s Minimax Hedging Algorithm

This section of the appendix provides an overview of Howe and Rustem’s minimax approach to portfolio hedging in the presence of transaction costs. Howe and Rustem (1997, pp. 1073-1075) discuss various hedging strategies that the writer of a call option can adopt to cover potential downside risk. The standard approach is delta hedging whereby the writer holds a number of units of the underlying stock so that any decrease in the value of the stocks is offset by an increase in the value of the option, and vice versa. The amount of underlying stock is given by the delta—the instantaneous derivative of the option price with respect to the actual stock price—that will change with time. When transactions costs are significant, the writer cannot engage in continuous rebalancing and must, instead, fall back on discrete delta hedging, rebalancing at discrete intervals of time. This results in the accumulation of hedging errors *HE*, over each of the intervals from *t* to *t*+1 as in:

$$HE = N\left(B_t - B_{t+1}\right) + n_t\left(S_{t+1} - S_t\right),$$

where the contract is initiated at time 0, $B = B(S, t)$ is the call price, S is the stock price at time t, n is the number of shares to hold, and N is the contracted number of shares of stock. The larger the hedging error the larger the cost of rebalancing. Howe and Rustem (p. 1075) observe that minimax control can be applied in the case of discrete delta hedging with the objective of finding the hedge ratio that minimizes the worst case hedging error. They follow Leland's approach to the problem (1985), that introduces a modification σ' to the volatility σ appearing in the original Black and Scholes formula as in the expression:

$$\sigma' = \sqrt{\sigma^2\left[1 + \sqrt{\frac{2}{\pi}}\frac{K}{\sigma\sqrt{\Delta t}}\right]},$$

where K is the roundtrip transaction cost arising from a buy and a sell of the same security, in either order. A formal representation of the problem is given by:

$$\min_{x \in R^n} \max_{y \in Y} f(x, y),$$

where y is a convex and compact infinite set $Y \subset \mathbb{R}^m$, f: $\mathbb{R}^n \times Y \Rightarrow \mathbb{R}^1$ such that f and $\nabla_x f(x, y)$ are continuous on $\mathbb{R}^n \times Y$. Constrained nonlinear programming cannot be used here, because it is a semi-finite optimization problem with an infinite number of constraints, each corresponding to the infinite number of elements in the set Y. In its place, Howe and Rustem present a quasi-Newton algorithm based on a quadratic approximation to the original problem. They consider a one-period and two-period application of the algorithm to both individual options and portfolios of options. In these applications, upper and lower bounds are imposed on the underlying asset price for each period and the worst-case hedging cost is determined within these bounds.

Minimum Martingale Measures

A recent development in the option pricing literature involves the application of minimum martingale measures (MMM) to the pricing of options in markets that feature volatility and transaction clustering. Typically, options are priced over a binomial lattice characterized by a mixed point process with constant jump sizes a, occurring over random or irregular rather than fixed time arrival intervals, and with the jump probabilities of up-moves and down-moves given by a logistic transformation of an autoregressive process (see Prigent, Renault and Scaillet, 1999a).

These marked point process (MPP) models allow for the possibility of either herding behaviour or mean reversion in the stock price and capture the volatility smiles and smirks that are observed in actively traded markets. They are directly analogous to the GARCH and log-GARCH models but are applied to durations. Previous research by the same group of authors (1999b) led them to prefer the log form of the autoregressive conditional duration (ACD) model of Engle and Russell (1998) over the unlogged form (to guarantee positivity of the durations), over a latent geometric brownian motion process, observable only when its logarithm crossed boundaries spaced by the pregiven jump size a (which was difficult to express in kernel form), and over a Poisson version (Bossaerts et al., 1996) of the MPP (because exponentially distributed inter-trade durations were not supported by the data for small values of a). The relevant durations are a product of two terms: a residual and the conditional expectation of the duration. The distribution of the residual, which can be determined from the data, can take either the exponential or Weibull form.

The conditional distribution of marks was represented by the logistic linear model of Cox (1981) extended to incorporate lagged values of the conditional probabilities: a binomial (ACB) version of the autoregressive conditional multinomial (ACM) model of Russell and Engle (1998). This econometric model possesses the obvious advantage that it can be estimated from actual high-frequency transaction data.

The presence of jumps in the model implies that the market is incomplete so that a method must be chosen to select one measure from amongst the class of equivalent martingale measures. Prigent, Renault and Scaillet, (1999a) reject the optimal variance measure that minimizes total risk from time t to T, under historical probabilities, because its existence cannot be guaranteed and, in general, it does not possess an analytical form. Instead, they choose Schweizer's (1991) approach based on minimizing local risk over successive small periods between time t and T. This measure is characterized by the fact that it sets to zero all risk premia on sources of risk orthogonal to the martingale part of the underlying's price process. An explicit form for the Radon-Nikodym derivative can always be constructed for this particular measure, which possess good convergence properties (Musiela and Rutkowski, section 4.2, pp. 99-108 and sections 10.2.2.-10.2.3. pp. 252-264, 1997). Moreover, jump boundedness ensures that the MMM is always positive so that the value of the trading strategy is an actual non-arbitrage price. Notably, when the mean-variance tradeoff (i.e. market price of risk) is deterministic the MMM is the closest of all EMMs to the original probability measure when measured by the relative entropy or directed divergence criterion (Föllmer and Schweizer, 1991).

However, this literature is less relevant to my application of option pricing theory in this paper, which examines real options under uncertainty aversion. For this purpose an approach based solely on minimum cross entropy methods is completely adequate.