

Teaching Notes #5
Ambiguity Aversion and Robust Decisions
Making¹

Pietro Veronesi
Graduate School of Business
University of Chicago
Business 35909 Spring 2005

© by Pietro Veronesi

This Version: November 17, 2005

¹These teaching notes draw heavily on the papers quoted in the references or in the Syllabus. They are intended for students of Business 537 only. Please, do not distribute without my prior consent.

1 Introduction

- In these teaching notes we will discuss the concept of ambiguity (or Knightian uncertainty) and Robust decision making.
- These are very recent developments in decision theory and optimal control theory to give a role for “ambiguous” events (examples will follow) and differentiate the concept of “risk aversion” from the concept of “ambiguity aversion” (or aversion to model misspecification).
- Early papers are Dow and Werlang (1992) “Uncertainty Aversion, Risk Aversion and the Optimal Choice of Portfolio), *Econometrica*, and Epstein and Wang (1994), “Intertemporal Asset Pricing under Knightian Uncertainty” *Econometrica*.
- We will discuss the following papers:
 1. Chen and Epstein (2002) “Ambiguity, Risk and Asset Returns in Continuous Time,” *Econometrica*
 2. Anderson, Hansen and Sargent (2000) “Robustness, Detection and the Price of Risk”
 3. Hansen, Sargent, Turmuhambetova and Williams (2001) “Robustness and Uncertainty Aversion”
- We start with Chen and Epstein (2002) which very close to the set up we already used in previous teaching notes.
- Hansen, Sargent, Turmuhambetova and Williams (2001) show the differences – and similarities – between the robust decision making approach of Anderson, Hansen and Sargent (2000) and the setting in Chen and Epstein (2002)).

- We will finally look at additional portfolio and asset pricing implications of robust choice using the set up of Maenhout (2004).

2 Knighthian Uncertainty

- Ellsberg Paradox:
 - There are two urns with red and black balls. Urn I has known composition (say 50-50), while Urn II has unknown composition.
 - A decision maker is confronted with the following bets:
 1. (R_I) Bet \$10 on Red from Urn I;
 2. (B_I) Bet \$10 on Black from Urn I;
 3. (R_{II}) Bet \$10 on Red from Urn II
 4. (B_{II}) Bet \$10 on Black from Urn II
 - It is often found the following pattern (with minor variations across experiments) of preferences

$$R_I \sim B_I \tag{1}$$

$$R_{II} \sim B_{II} \tag{2}$$

$$R_I \succeq R_{II} \tag{3}$$

$$B_I \succeq B_{II} \tag{4}$$

- Clearly, these preferences are inconsistent with probability theory.
- In fact, if we call p the probability to extract a red ball from urn II , from (3) one could conclude that $p < .5$ while from (4) one could conclude that $p > .5$.

- In addition, (2) seem to indicate $p = .5$.
- Bets such as R_I and B_I are called “roulette lotteries” in the sense that the odds are known (it is like spinning a roulette);
- Bets such as R_{II} and B_{II} are sometimes called “horse lotteries” in the sense that the odds are not known, but agents may have a probability distribution on them.
- Lotteries on the first urn characterize “Risk” (in the common sense);
- Lotteries on the second urn characterize what is called (Knightian) “uncertainty” or “ambiguity.”

2.1 Multiple Priors to “solve” the Paradox

- How do we make decisions in the presence of ambiguity (or knightian) uncertainty?
- Somebody could be a pure Bayesian and do not get trapped in the tricks of the thought experiment (1)-(4).
- Indeed, Savage axioms have been put forward most often (and by himself) as *normative*! That is, how to guide reasonable choices given a set of “axioms” that make sense ex-ante.
- In the case of the experiment (1) - (4), Savage would say: Given your preferences established in (1) and (2), you must be also indifferent between R_I and R_{II} and B_I and B_{II} .
- On the other hand, many people even after reading Savage axioms, agreeing that they are sensible, when faced with bets 1 to 4, they would still have preferences as in (1) - (4).

- Other researchers have turned to other axioms that would be compatible with those preferences.
- Typical two references are the original Schmeidler (1989) and Gilboa and Schmeidler (1993).
- Depending on tastes, one can view Schmeidler (1989) and Gilboa and Schmeidler (1993) as more or less reasonable than Savage's (or Anscombe and Auman (1963)'s) axioms.
- Rather than delving in issues on decision theory, we give the “answer” to the basic question on how to incorporate the above preferences in an expected utility type of framework.
- It turns out that if the decision maker has a set of priors on the probability of getting “red” from urn II and he/she uses the max-min rule, one can “resolve” the paradox.
- Example, suppose that the subjective distribution on extracting red from urn II is $p \in [.4, .6]$;
- Then, the max-min rule applied to (3) would imply that the agent uses the “worst” case scenario for “ R_{II} ” and hence use $p = .4$. In this case, R_I would indeed be preferred to R_{II} ;
- Similarly, in (4) the worst case scenario for B_{II} is that $p = .6$, which again implies that agent prefers B_I to B_{II} .
- However, the symmetry in the possible priors also imply that (2) is satisfied.
- So, in what follows we shall assume that agents are endowed with a class of probability distributions \mathcal{P} on the event on the state of nature Ω ;

- Then, they take decisions according to the rule

$$\max_c \min_{P \in \mathcal{P}} E^P [u(c)]$$

3 Knighthian Uncertainty in Continuous Time

- Consider a probability space (Ω, \mathcal{F}, P) on which is defined a standard d -dimensional Brownian motion \mathbf{B} along with the natural filtration $\{\mathcal{F}_t\}$.
- As usual, the filtration $\{F_t\}$ is augmented by the P -null sets.
- Consumption processes c are defined in the usual square integrable space L_+^2 .
- As in TN4, given a consumption process c , consider the Stochastic Differential Utility (SDU)

$$V_t^P = E_t^P \left[\int_t^T f(c_\tau, V_\tau^P) d\tau \right]$$

- where $f(.,.)$ is the *normalized aggregator* defined in TN4 (and Duffie and Epstein);
- Recall that alternatively, V_t^P is characterized as being the unique solution to the Backward Stochastic Differential Equation

$$dV_t^P = -f(c_t, V_t^P) dt + \boldsymbol{\sigma}_t d\mathbf{B}_t \quad (5)$$

- with $V_T^P = 0$ and for a volatility process $\boldsymbol{\sigma}_t$ which is endogenously determined.

- The superscript P denotes the fact that the value of the SDU is computed using the probability P as reference probability.
- So far, there is no “ambiguity”: All the expectations are taken with respect to the reference probability P , which, in a rational expectations paradigm, would be considered the “true” probability distribution.
- The difficulty is to define a class of probability distributions \mathcal{P} in a convenient way.
- Chen and Epstein (2002) consider “density generators”

3.1 Density Generators

- The idea is to construct a set of probabilities Q that are equivalent to P .
- As one can imagine, in continuous time this is possible through the use of Novikov’s theorem and Girsanov’s theorem.
- Indeed, a density generator is a \mathcal{R}^d -valued process $\boldsymbol{\theta}$ satisfying the Novikov’s condition

$$E \left[\exp \left(\frac{1}{2} \int_0^T \boldsymbol{\theta}'_{\tau} \boldsymbol{\theta}_{\tau} d\tau \right) \right] < \infty$$

- so that the process

$$z_t^{\boldsymbol{\theta}} = \exp \left(-\frac{1}{2} \int_0^t \boldsymbol{\theta}'_{\tau} \boldsymbol{\theta}_{\tau} d\tau - \int_0^t \boldsymbol{\theta}'_{\tau} d\mathbf{B}_{\tau} \right)$$

- is a P -martingale.

- As we have seen in TN1, we have that z_T^θ defines a probability measure Q^θ that is equivalent to P on Ω through the Radon-Nikodym derivative

$$\frac{dQ^\theta}{dP} = z_T^\theta$$

- Hence, we can identify a class of probability measures \mathcal{P} (equivalent to each other) with a set Θ of density generators.
- Chen and Epstein (2002) define the set of density generators by starting from a set $\{\Theta_t\}$ of *correspondences*² from Ω to \mathcal{R}^d , with a number of characteristics:

1. **Uniform boundedness:** There exists a set $\mathcal{K} \subset \mathcal{R}^d$ such that $\Theta_t(\omega) \in \mathcal{K}$ for all $\omega \in \Omega$ and $t \in [0, T]$;
2. **Compact-Convex:** The set Θ_t is compact and convex valued;
3. **Measurability:** The correspondence $(t, \omega) \rightarrow \Theta_t(\omega)$ is measurable with respect to the product sigma-algebra $\mathcal{B}([0, s]) \times \mathcal{F}_s$ for any $0 < t \leq s \leq T$.
4. **Normalization:** $\mathbf{0} \in \Theta_t(\omega)$ $\ell \times P$ -almost surely (where ℓ is the Lebesgue measure on $[0, T]$).

- Clearly, 1 is needed to make sure that z_t^θ is a martingale. Condition 2 will be needed to reach a minimum point in the max-min rule, 3 just requires that we can put probability measures on the events generated by the correspondence and

²A correspondence is a generalization of a function: It is a map that associates to each element in the domain $(t, \omega) \in [0, T] \times \Omega$ a set $\Theta_t(\omega) \subseteq \mathcal{R}^d$ rather than a point in \mathcal{R}^d .

4 just makes sure that the original probability measure P is included in the set \mathcal{P} that we will construct.

- Hence, the set of density generators is defined by

$$\Theta = \{ \{ \boldsymbol{\theta}_t \} : \boldsymbol{\theta}_t(\omega) \in \Theta_t(\omega) \ \ell \times P - \text{a.s.} \} \quad (6)$$

- The form of Θ restricts the possible probability measures but is quite general.
- Specifically, it requires the probability across time and states to be “rectangular,” that is, obtained by the multiplication of their projections (marginals) on the two dimensions..
- We need one more definition: We say that Θ is *stochastically convex* if for any real valued process $\{ \lambda_t \}$ with $0 \leq \lambda_t \leq 1$ we have

$$\boldsymbol{\theta}, \boldsymbol{\theta}' \in \Theta \implies \{ \lambda_t \boldsymbol{\theta}_t + (1 - \lambda_t) \boldsymbol{\theta}'_t \} \in \Theta$$

- It turns out that set of density generators Θ defined in (6) has a number of nice properties.
- In particular, it is bounded and stochastically convex.
- In addition, for any \mathcal{R}^d -valued process $\{ \sigma_t \}$, there exists a process $\{ \boldsymbol{\theta}_t^* \} \in \Theta$ such that

$$\boldsymbol{\theta}_t^* \cdot \sigma_t = \max_{\boldsymbol{\theta}_t \in \Theta_t} \boldsymbol{\theta}_t \cdot \sigma_t$$

3.2 The Class of Priors

- Given the density generators Θ , we can define the class of prior distributions on Ω “in the obvious way” as follows:

$$\mathcal{P}^\Theta = \left\{ Q^\theta : \theta \in \Theta \text{ and } \frac{dQ^\theta}{dP} = z_T^\theta \right\}$$

- Given the nice properties of Θ it turns out that also \mathcal{P}^Θ has nice properties that allow one to take minimizations across the various probabilities $Q \in \mathcal{P}^\Theta$ in a meaningful way.
- In fact, $P \in \mathcal{P}^\Theta$ and each $Q \in \mathcal{P}^\Theta$ is equivalent to P .
- Also, it is convex, compact (in some sense) and, in addition, has the following important property
- For every $X \in \mathcal{L}^2(\Omega, \mathcal{F}_T, P)$, there exists $Q^* \in \mathcal{P}^\Theta$ such that

$$E^{Q^*} [X | \mathcal{F}_t] = \min_{Q \in \mathcal{P}^\Theta} E^Q [X | \mathcal{F}_t] \text{ for all } 0 \leq t \leq T$$

- That is, we can select a probability measure Q^* from the class of priors \mathcal{P}^Θ such that a minimum conditional expectation with respect to any probability measure is attained.

3.3 The Main Result: Multiple-Prior Recursive Utility

- Consider aggregators $f : C \times \mathcal{R} \rightarrow \mathcal{R}$ that satisfy
 - Borel measurability

– Uniform Lipschitz: That is, there exists k such that

$$|f(c, v) - f(c, w)| \leq k|v - w| \text{ for all } c \in C \text{ and } v, w \in \mathcal{R}$$

– Growth condition $E \left[\int_0^T f^2(c_t, w) dt \right] < \infty$ for all consumption c in the domain D .

- The main result of Chen and Epstein (2002) is as follows:
- **Theorem:** Let the set of density generators Θ and the aggregator $f(., .)$ satisfy the conditions above.
- Fix a consumption process c . Then

– (a) There exists a unique (continuous) process $\{V_t\}$ satisfying the (Backward) Stochastic Differential Equation

$$dV_t = \left[-f(c_t, V_t) + \max_{\theta \in \Theta} \theta_t \cdot \sigma_t \right] dt + \sigma_t \cdot d\mathbf{B}_t \quad (7)$$

– with $V_T = 0$;

– (b) For each $Q = Q^\theta \in \mathcal{P}^\Theta$, denote by (V_t^Q) the unique solution to

$$dV_t^Q = \left[-f(c_t, V_t^Q) + \theta_t \cdot \sigma_t^Q \right] dt + \sigma_t^Q \cdot d\mathbf{B}_t \quad (8)$$

– with $V_T^Q = 0$. Then $\{V_t\}$ defined in (a) is the unique solution to

$$V_t = \min_{Q \in \mathcal{P}^\Theta} V_t^Q \quad (9)$$

- and there exists $Q^{\theta^*} \in \mathcal{P}^\Theta$ such that $V_t = V_t^{Q^*}$, $0 \leq t \leq T$.
- (c) The process $\{V_t\}$ is the unique solution to $V_T = 0$ and

$$V_t = \min_{Q \in \mathcal{P}^\Theta} E^Q \left[\int_t^\tau f(c_s, V_s) ds + V_\tau | \mathcal{F}_t \right], \quad 0 \leq t < \tau \leq T \quad (10)$$

- Part (a) and equation (7) show the clear relationship between the stochastic differential utility framework in (5) and the *multiple prior, recursive utility*.
- To understand it better, consider again the standard Stochastic Differential Utility as in (5) with respect to any of the probabilities $Q = Q^\theta \in \mathcal{P}^\Theta$. By definition, we must have

$$dV_t^Q = -f(c_t, V_t^Q) dt + \boldsymbol{\sigma}_t \cdot d\mathbf{B}_t^Q \quad \text{and} \quad V_T^Q = 0 \quad (11)$$

- where \mathbf{B}^Q is a Brownian motion under Q .
- By Girsanov's theorem, we can re-express (11) in terms of the original Brownian motion under P , since we know that $\mathbf{B}_t^Q = \int_0^t \boldsymbol{\theta}'_\tau d\tau + \mathbf{B}_t$

$$dV_t^Q = \left[-f(c_t, V_t^Q) + \boldsymbol{\theta}'_t \cdot \boldsymbol{\sigma}_t \right] dt + \boldsymbol{\sigma}_t \cdot d\mathbf{B}_t \quad \text{and} \quad V_T^Q = 0 \quad (12)$$

- Comparing (12) with (7) we see that the difference lies in the maximization over $\boldsymbol{\theta}_t \cdot \boldsymbol{\sigma}_t$ at any instant. This is a penalty imposed on the process for the utility function, where its

value depends on the set of possible prior distributions \mathcal{P}^Θ , or equivalently Θ .

- Why $\max_{\theta \in \Theta} \theta_t \cdot \sigma_t$ is a penalty? (It looks like we are maximizing!)
- The reason is that recall that the equation is a *Backward* SDE, that is, we must arrive at $V_T = 0$.
- If we have a higher drift in the process for V_t (as is the case if we take the maximum), intuitively we must start from a lower value of V_t in order to reach $V_T = 0$ at maturity (recall that the consumption c is fixed!)
- This is confirmed by the result in part (b): We indeed have that the multiple-prior, recursive utility V_t is the minimum of all the SDU computed across the various probability measures.
- This shows that the agent takes the “worst case scenario” as the proper scenario.
- We also notice that “*ambiguity*” in this framework is therefore limited to the *drift* of the driving process.
- Finally, part (c) show that the multiple prior, recursive utility $\{V_t\}$ is indeed recursive, and hence, time consistent.
- One should notice that this result is not immediate as in the case of time-separable utilities, as we show below.

3.4 Example: Time Separable Preferences

- Consider the time-separable utility function.

- As shown in TN4, this correspond to the case

$$f(c, v) = u(c) - \phi v$$

- In fact, in this case we showed that for each $Q = Q^\theta \in \mathcal{P}^\Theta$

$$V_t^Q = E^Q \left[\int_t^T e^{-\phi(s-t)} u(c_s) ds | \mathcal{F}_t \right]$$

- is the solution to (11). By Girsanov's theorem, we can represent (11) as (8) in part (b) of the Theorem.
- Hence, we can apply the theorem, so that we have the closed form solution

$$V_t = \min_{Q \in \mathcal{P}^\Theta} E^Q \left[\int_t^T e^{-\phi(s-t)} u(c_s) ds | \mathcal{F}_t \right] \quad (13)$$

- Notice that *a-priori* it is not obvious that V_t defined in this way satisfies dynamic consistency.
- Actually, in general it is not true.
- In fact, in the usual case (without multiple priors), one can use the linearity of the expectation operator to write

$$\begin{aligned} V_t &= E \left[\int_t^T e^{-\phi(s-t)} u(c_s) ds | \mathcal{F}_t \right] \\ &= E \left[\int_t^\tau e^{-\phi(s-t)} u(c_s) ds + e^{-\phi(\tau-t)} V_\tau | \mathcal{F}_t \right] \end{aligned}$$

- From here, dynamic consistency is immediate.

- In case (13) instead there is also a minimum operator in the middle, which is *not* a linear operator.
- However, the theorem shows dynamic consistency in (10). That is, we indeed have

$$V_t = \min_{Q \in \mathcal{P}^\Theta} E^Q \left[\int_t^\tau e^{-\phi(s-t)} u(c_s) ds + e^{-\phi(s-t)} V_\tau | \mathcal{F}_t \right]$$

- It turns out that the assumption on Θ is important in determining this result.

4 Optimal Consumption

- We restrict the analysis to the case of time-separable utility functions (for now).
- Consider a standard asset market. There are d securities with prices \mathbf{S}_t evolving according to the Ito process

$$d\mathbf{S}_t = \mathbf{I}_S \boldsymbol{\mu}_{S,t} dt + \mathbf{I}_S \boldsymbol{\sigma}_{S,t} d\mathbf{B}_t$$

- where \mathbf{B} is a d -dimensional Brownian motion. The processes $\boldsymbol{\mu}_S$ and $\boldsymbol{\sigma}_S$ satisfy the usual restrictions.
- There is a riskless asset, with instantaneous risk-free rate of r_t and price

$$\beta_t = \beta_0 e^{\int_0^t r_s ds}$$

- Assume market completeness, so that we can define a process

$$\boldsymbol{\nu}_t = \boldsymbol{\sigma}_{S,t}^{-1} \cdot (\boldsymbol{\mu}_{S,t} - r_t \mathbf{1}_d)$$

- This is the standard market price of risk process. Chen and Epstein (2002) call it “*market price of uncertainty*” because its size (in equilibrium) depends both on “risk” and “ambiguity,” as we shall see.
- The other definitions are standard: Let c denote a consumption process in L_+^2 and let $(\varphi^0, \boldsymbol{\varphi})$ denote a trading strategy, where φ^i is the fraction of wealth invested in asset i .
- Assume w is the investor’s initial wealth, so that the dynamic budget constraint is

$$W_t = w + \int_0^t W_\tau \left[r_\tau + \boldsymbol{\varphi}_\tau \cdot (\boldsymbol{\mu}_{S,\tau} - r_t \mathbf{1}_d) \right] - c_\tau d\tau + \int_0^t W_\tau \boldsymbol{\varphi}_\tau \boldsymbol{\sigma}_{S,\tau} d\mathbf{B}_\tau \quad (14)$$

- The problem of the agent is then to solve

$$\max_{c, \boldsymbol{\varphi}} \min_{Q \in \mathcal{P}^\Theta} E^Q \left[\int_0^T e^{-\phi t} u(c_t) dt \right] \quad (15)$$

- subject to the dynamic budget constraint (14).
- Before moving to the effective methodology to solve for consumption, we notice that for each $Q = Q^\theta \in \mathcal{P}^\Theta$, we can rewrite the process for stock returns under the new probability measure Q as

$$\begin{aligned} d\mathbf{S}_t &= \mathbf{I}_S \boldsymbol{\mu}_{S,t} dt + \mathbf{I}_S \boldsymbol{\sigma}_{S,t} d\mathbf{B}_t \\ &= \mathbf{I}_S \boldsymbol{\mu}_{S,t} dt + \mathbf{I}_S \boldsymbol{\sigma}_{S,t} (d\mathbf{B}_t^Q + \boldsymbol{\theta}_t dt) \\ &= (\mathbf{I}_S \boldsymbol{\mu}_{S,t} + \boldsymbol{\theta}_t) dt + \mathbf{I}_S \boldsymbol{\sigma}_{S,t} d\mathbf{B}_t^Q \end{aligned}$$

- Hence, the class of priors \mathcal{P}^Θ define a set of processes for stock returns, which differ in the drift only.
- The agent will take decisions based on a “worst case” scenario. Since intuitively this “worst case scenario” will tend to consider returns as lower than what they in fact are, this will induce a lower demand and hence, in equilibrium, a higher risk premium and lower interest rate.
- On the other hand, other issues come up, as we will see.
- Consider the state-price density

$$\chi_t = \exp \left(- \int_0^t \left(r_\tau + \frac{1}{2} \boldsymbol{\nu}_\tau \boldsymbol{\nu}'_\tau \right) d\tau - \int_0^t \boldsymbol{\nu}_\tau d\mathbf{B}_\tau \right)$$

- We then know that market completeness implies that we can transform the dynamic budget constraint (14) into the static one

$$E^P \left[\int_0^T \chi_t c_t \right] \leq w \quad (16)$$

- so that the problem reduces to

$$\max_c \min_{Q \in \mathcal{P}^\Theta} E^Q \left[\int_0^T e^{-\phi t} u(c_t) dt \right] \quad (17)$$

- subject to (16). In fact, market completeness ensures that once we solved for consumption, we can always find the optimal trading strategy φ that finances it.

- From the Theorem in the previous section, we know that for every c , we can find a d -valued process $\theta_c^* \in \Theta$ such that

$$E^{\theta_c^*} \left[\int_0^T e^{-\phi t} u(c_t) dt \right] = \min_{Q \in \mathcal{P}^\Theta} E^Q \left[\int_0^T e^{-\phi t} u(c_t) dt \right]$$

- where $E^{\theta_c^*}$ denotes the expectation operator with respect to $Q^{\theta_c^*}$.
- However, we also know that by using the Radon-Nikodym theorem, we can also rewrite (a little review here:)

$$\begin{aligned} E^{\theta_c^*} \left[\int_0^T e^{-\phi t} u(c_t) dt \right] &= E^P \left[z_T^{\theta_c^*} \int_0^T e^{-\phi t} u(c_t) dt \right] \\ &= E^P \left[\int_0^T e^{-\phi t} u(c_t) z_T^{\theta_c^*} dt \right] \\ \text{(Law of It. Exp.)} &= E^P \left[E_t^P \left(\int_0^T e^{-\phi t} u(c_t) z_T^{\theta_c^*} dt \right) \right] \\ \text{(Fubini Theorem)} &= E^P \left[\int_0^T e^{-\phi t} E_t^P \left(u(c_t) z_T^{\theta_c^*} \right) dt \right] \\ \text{(\mathit{c}_t \mathcal{F}_t\text{-measurable})} &= E^P \left[\int_0^T e^{-\phi t} u(c_t) E_t^P \left(z_T^{\theta_c^*} \right) dt \right] \\ \text{(\mathit{z}_t^{\theta_c^*} \text{ is a martingale})} &= E^P \left[\int_0^T e^{-\phi t} u(c_t) z_t^{\theta_c^*} dt \right] \end{aligned}$$

- Hence, we finally have the following equivalent problem

$$\max_c E^P \left[\int_0^T e^{-\phi t} u(c_t) z_t^{\theta_c^*} dt \right] \tag{18}$$

- subject to

$$E^P \left[\int_0^T \chi_t c_t \right] \leq w \tag{19}$$

- By using a Lagrange multiplier method, a necessary condition is then

$$e^{-\phi t} u_c(c_t) z_t^{\theta_c^*} = \lambda \chi_t \quad (20)$$

- for some lagrange multiplier λ ; or

$$c_t^* = \mathcal{I}_u \left(\lambda^* \frac{\chi_t}{z_t^{\theta_c^*}} e^{\phi t} \right)$$

- where \mathcal{I}_u is the inverse of the marginal utility of consumption and λ^* is determined by the budget constraint

$$E^P \left[\int_0^T \chi_t \mathcal{I}_u \left(\frac{\lambda^* \chi_t e^{\phi t}}{z_t^{\theta_c^*}} \right) \right] = w \quad (21)$$

- The multiple priors are showing themselves in the first order conditions (20) through the (martingale) process $z_t^{\theta_c^*}$.
- Recall that in the case of a singular prior we indeed have $z_t^{\theta_c^*} = 1$ and hence we obtain the usual solution for optimal consumption.
- Multiple priors affect the condition for optimal consumption by adding a multiplicative process to the marginal utility itself.
- One can expect that this extra-process may affect the equilibrium interest rate and equity price because it may increase or decrease the variation in the equilibrium marginal utility of consumption.

- Indeed, the effect of multiple priors is to obtain a new state-price density

$$\chi_t^* = \frac{\chi_t}{z_t^{\theta_c^*}}$$

- which modify consumption (and hence the trading strategy):

$$c_t^* = \mathcal{I}_u \left(\lambda \chi_t^* e^{\phi t} \right)$$

4.1 Implication for Interest Rates and Stock Returns

- We now apply the same methodology used in TN2 to find the equilibrium interest rate and expected risk premia.
- We “turn around” condition (20) and impose that in equilibrium it has to hold for a representative agent, at the point where $c_t = e_t$ (where e_t is an endowment process, or the dividend of the security)

$$e^{-\phi t} u_c(e_t) z_t^{\theta_c^*} = \lambda \chi_t \tag{22}$$

- This equilibrium condition makes clear that depending on whether $z_t^{\theta_c^*}$ and $u_c(e_t)$ are positively or negatively correlated, interest rates and premia will vary.
- In fact, if they are positively correlated, then periods of high marginal utility (bad periods) also have a high weight on the marginal utility, which make the “bad period” feel even worse. That is, the state price χ_t must increase even more, making this state more expensive.

- If $z_t^{\theta^*}$ and $u_c(e_t)$ are negatively correlated, bad and good periods are characterized by similar state prices χ_t because, in a sense, the existence of multiple priors work as a natural hedge.
- This implies that there is not going to be a premium for ambiguity, but a discount: Investors are happy to be averse to uncertainty.

- We now check this more formally:

- Let the aggregate endowment evolve according to the Ito process

$$\frac{de_t}{e_t} = \mu_{e,t}dt + \boldsymbol{\sigma}_{e,t}d\mathbf{B}_t$$

- and recall that by definition of z_t^θ (suppress the index * and c) we have

$$z_t^\theta = \exp\left(-\frac{1}{2}\int_0^t \boldsymbol{\theta}'_\tau \boldsymbol{\theta}_\tau d\tau - \int_0^t \boldsymbol{\theta}'_\tau d\mathbf{B}_\tau\right)$$

- that is,

$$dz_t^\theta = -z_t^\theta \boldsymbol{\theta}_t d\mathbf{B}_t$$

- Define $F(t, e_t, z_t^\theta) = e^{-\phi t} u_c(e_t) z_t^\theta$, so that by Ito's lemma we have

$$\begin{aligned} dF &= F_t dt + F_e de_t + F_z dz_t^\theta + \frac{1}{2} F_{ee} de_t^2 + \frac{1}{2} F_{zz} (dz_t^\theta)^2 + F_{ez} de_t dz_t^\theta \\ &= -\phi F dt + \frac{u_{cc}(e_t)}{u_c(e_t)} F de_t + F \frac{dz_t^\theta}{z_t^\theta} + \frac{1}{2} \frac{u_{ccc}(e_t)}{u_c(e_t)} F de_t^2 \end{aligned}$$

$$+ \frac{u_{cc}(e_t)}{u_c(e_t)} F \frac{dz_t^\theta}{z_t^\theta} de_t$$

- Hence

$$\frac{dF}{F} = - \left(\phi + \gamma(e_t) \mu_{e,t} - \frac{1}{2} \gamma(e_t) q(e_t) \boldsymbol{\sigma}_{e,t} \boldsymbol{\sigma}'_{e,t} - \gamma(e_t) \boldsymbol{\theta}_t \boldsymbol{\sigma}'_{e,t} \right) dt - (\gamma(e_t) \boldsymbol{\sigma}_{e,t} + \boldsymbol{\theta}_t) d\mathbf{B}_t$$

- where

$$\gamma(e_t) = - \frac{e_t u_{cc}(e_t)}{u_c(e_t)}$$

$$q(e_t) = - \frac{e_t u_{ccc}(e_t)}{u_{cc}(e_t)}$$

- The right hand side of (22) is given instead by

$$\frac{d\chi_t}{\chi_t} = -r_t dt - \boldsymbol{\nu}_t d\mathbf{B}_t$$

- Hence, equating as usual the drifts we find

$$r_t = \phi + \gamma(e_t) \mu_{e,t} - \frac{1}{2} \gamma(e_t) q(e_t) \boldsymbol{\sigma}_{e,t} \boldsymbol{\sigma}'_{e,t} - \gamma(e_t) \boldsymbol{\theta}_t \boldsymbol{\sigma}'_{e,t}$$

- We see that one more term appears in the risk-free rate: This depends on the relative risk aversion coefficient $\gamma(e_t)$ as well as

$$-\boldsymbol{\theta}_t \boldsymbol{\sigma}'_{e,t} = cov_t \left(\frac{dz_t^\theta}{z_t^\theta}, \frac{de_t}{e_t} \right)$$

- As previous intuition suggested, when $cov_t \left(dz_t^\theta / z_t^\theta, de_t / e_t \right) < 0$, then multiple priors are “bad,” hence agents increase the demand for bonds, decreasing the equilibrium interest rate.
- Turning to risk premia, by equating the diffusions, we find the excess returns

$$\boldsymbol{\mu}_{S,t} - r_t \mathbf{1}_d = \gamma(e_t) \boldsymbol{\sigma}_{S,t} \boldsymbol{\sigma}'_{e,t} + \boldsymbol{\sigma}_{S,t} \boldsymbol{\theta}'_t$$

- or

$$E_t \left[\frac{dS_t^i}{S_t} \right] - r_t = \gamma(e_t) cov_t \left(\frac{de_t}{e_t}, \frac{dS_t^i}{S_t^i} \right) - cov_t \left(\frac{dz_t^\theta}{z_t^\theta}, \frac{dS_t^i}{S_t^i} \right)$$

- The first term is the standard C-CAPM. The second term shows that the deviation from C-CAPM implied by the multiple prior setting depend on the covariance of each asset with z_t^θ .
- Again, recall that z_t^θ impacts the equilibrium level of the marginal utility

$$e^{-\phi t} u_c(e_t) z_t^\theta = \lambda \chi_t \tag{23}$$

- Hence, those asset that have a positive covariance with z_t^θ are a natural hedge against changes in z_t^θ , because they make increase the state price density χ_t exactly when also z_t^θ increases, stabilizing the marginal utility (and hence consumption).

4.2 Kreps-Porteus/Epstein-Zin Preferences

- Similar results obtain under the Kreps-Porteus, Epstein-Zin preference.
- In this case we have

$$f(c, v) = \frac{c^\rho - \phi(\alpha v)^{\rho/\alpha}}{\rho(\alpha v)^{(\rho-\alpha)/\alpha}}$$

- where recall that the degree of intertemporal substitution is $\psi = (1 - \rho)^{-1}$ and the coefficient of risk aversion is α .
- By using “supergradients,” as in Duffie and Skiadas (1994) and Schroder and Skiadas (1999), Chen and Epstein (2002) find that the First Order Condition is

$$\exp\left(\int_0^t f_v(c_s, V_s(c)) ds\right) f_c(c_t, V_t(c)) z_t^{\theta^*} = f_c(c_0, V_0) \chi_t \quad (24)$$

- Notice that by setting $z_t^{\theta^*} = 1$, this generalizes the result in TN4 about Stochastic Differential Utility. Such a result has been found by Schroeder and Skiadas (1999).
- By using a similar strategy as above (i.e. apply Ito’s Lemma to (24) and matching drift and diffusion of this with the state price density), Chen and Epstein (2002) find that the interest rate is given by

$$r_t = \phi + \frac{1}{\psi} \mu_{e,t} - \frac{1}{2} \frac{2 - \rho}{\psi} \sigma_{e,t} \sigma'_{e,t} - \frac{1}{2\psi} \theta_t \sigma'_{e,t} + \frac{(\alpha - \rho)}{\rho} \left(\bar{\sigma}_t \sigma'_{e,t} + \frac{1}{2} \psi \bar{\sigma}_t \bar{\sigma}'_t \right)$$

- where

$$\bar{\sigma}_t = \frac{\rho \sigma_{V,t}}{\alpha V_t} = \left(\sigma_{M,t} + \frac{1}{\psi} \sigma_{e,t} \right)$$

- In addition, the risk premium on the various securities is given by

$$\boldsymbol{\mu}_{S,t} - r_t \mathbf{1}_d = \frac{\alpha}{\psi \rho} \boldsymbol{\sigma}_{S,t} \boldsymbol{\sigma}'_{e,t} + \left(1 - \frac{\alpha}{\rho} \right) \boldsymbol{\sigma}_{S,t} \boldsymbol{\sigma}'_{M,t} + \boldsymbol{\sigma}_{S,t} \boldsymbol{\theta}'_t$$

- The “add on” due to ambiguity is clearly the same as before, where we notice that it is intertemporal substitution that plays a role in the determination on the interest rate, and not risk-aversion.

5 Robust Control

- A concept related to ambiguity is the model uncertainty and robust control put forward by Hansen, Sargent and Tallarini (1999).
- Here we will follow Anderson, Hansen and Sargent (2000) and, even more closely, Hansen, Sargent, Turmuhabetova and Williams (2001).
- Some of results are instead contained in Maenhout (2004).
- So far, we have said little about the class \mathcal{P}^Θ of probability distributions that the agent may have.
- The idea of “robust control” is that an agent (a decision maker) want to take decision based on a model that is robust to (small) specification errors.
- We formulate the problem directly within the securities market model specified above.
- That is, fix a probability space (Ω, P, \mathcal{F}) on which we define a d -dimensional Brownian motion B_t and a (standard) P -augmented filtration $\{\mathcal{F}_t\}$
- As in TN4, in the set up of Duffie and Eptein (1992), let there be n state variables $\mathbf{X} = (X^1, \dots, X^n)$ evolving according to the Ito processes

$$d\mathbf{X}_t = \mathbf{b}(\mathbf{X}_t, t) dt + \mathbf{a}(\mathbf{X}_t, t) d\mathbf{B}_t$$

- where $\mathbf{b} : R^n \times [0, T] \rightarrow R^n$ and $\mathbf{a} : R^n \times [0, T] \rightarrow R^{n \times d}$ and \mathbf{B}_t is the d dimensional Brownian motion introduced earlier.

- There are d securities with prices \mathbf{S}_t evolving according to the Ito process

$$d\mathbf{S}_t = \mathbf{I}_S \boldsymbol{\mu}_{S,t} dt + \mathbf{I}_S \boldsymbol{\sigma}_{S,t} d\mathbf{B}_t \quad (25)$$

- where

$$\boldsymbol{\mu}_{S,t} = \boldsymbol{\mu}_S(\mathbf{X}_t) \text{ and } \boldsymbol{\sigma}_{S,t} = \boldsymbol{\sigma}_S(\mathbf{X}_t)$$

- There is a risk-free security with short-rate process $r_t = r(\mathbf{X}_t)$.
- We assume again market completeness, so that we can define a process

$$\boldsymbol{\nu}_t = \boldsymbol{\sigma}_{S,t}^{-1} \cdot (\boldsymbol{\mu}_{S,t} - r_t \mathbf{1}_d)$$

- Let c denote a consumption process in L_+^2 and let $(\varphi^0, \boldsymbol{\varphi})$ denote a trading strategy, where φ^i is the fraction of wealth invested in asset i .
- Assume w is the investor's initial wealth, so that the dynamic budget constraint is

$$dW_t = \mu_W(W_t, \mathbf{X}_t, \boldsymbol{\varphi}_t, c_t) dt + \sigma_W(W_t, \mathbf{X}_t, \boldsymbol{\varphi}_t, c_t) d\mathbf{B}_t \quad (26)$$

- where

$$\begin{aligned} \mu_W(W_t, \mathbf{X}_t, \boldsymbol{\varphi}_t, c_t) &= W_t \left[r_t + \boldsymbol{\varphi}_t \cdot (\boldsymbol{\mu}_{S,t} - r_t \mathbf{1}_d) \right] - c_t \\ \sigma_W(W_t, \mathbf{X}_t, \boldsymbol{\varphi}_t, c_t) &= W_t \boldsymbol{\varphi}_t \boldsymbol{\sigma}_{S,t} \end{aligned}$$

- The standard problem of the agent who is not concerned about robustness is then

$$\max_{c, \varphi} E^P \left[\int_0^T e^{-\phi t} u(c_t) dt \right] \quad (27)$$

- subject to the dynamic budget constraint (26).

5.1 Relative Entropy

- It is assumed that the decision maker views (25) as an approximation and takes into account alternative models that are difficult to distinguish statistically from (25).
- This class of alternative models is obtained by “perturbing” the probability P that governs the dynamics of (25).
- Given the continuous time framework and what we have already learnt in the previous section, it is intuitive that “perturbation” in this context means changing the probability measure to an equivalent Q .
- Using Girsanov’s theorem, we will basically perturb the model to replace \mathbf{B} in (25) by

$$\mathbf{B}_t = \widehat{\mathbf{B}}_t + \int_0^t \mathbf{h}_s ds \quad (28)$$

- Hence, we obtain the perturbed process for the state variables as

$$d\mathbf{X}_t = \mathbf{b}(\mathbf{X}_t, t) dt + \mathbf{a}(\mathbf{X}_t, t) \left(d\widehat{\mathbf{B}}_t + \mathbf{h}_t dt \right) \quad (29)$$

$$dW_t = \mu_W(W_t, \mathbf{X}_t, \varphi_t, c_t) dt + \sigma_W(W_t, \mathbf{X}_t, \varphi_t, c_t) \left(d\widehat{\mathbf{B}}_t + \mathbf{h}_t dt \right) \quad (30)$$

- In order to decide when \mathbf{h}_t is too large (so that statistical inference would detect a large difference from the approximating model), AHS use the concept of “Relative Entropy,” defined as follows:
- Consider a scalar stochastic process $\{g_t\}$, progressively measurable.
- Take the product space $\Omega^* = \Omega \times \mathcal{R}^+$ and let \mathcal{F}^* be the smallest sigma-algebra containing $\mathcal{F}_t \times \mathcal{B}([0, t])$.
- Finally, let P^* be the product probability measure obtained by $P \times M$, where M is *exponentially distributed* with density $\delta \exp(-\delta t)$.
- If we denote by E^* the expectation with respect to P^* , by construction we have

$$E^*(g) = \delta \int_0^\infty \exp(-\delta t) E_P(g_t) dt$$

- Now, let Q be a probability measure that is equivalent to P . A similar construction yields a probability measure $Q^* = Q \times M$ such that

$$\begin{aligned} E_Q^*(g) &= \delta \int_0^\infty \exp(-\delta t) E_Q(g_t) dt \\ &= \delta \int_0^\infty \exp(-\delta t) E_P(\xi_t g_t) dt \end{aligned}$$

- where

$$\xi_T = \frac{dQ}{dP} = \exp\left(\int_0^T \mathbf{h}_t d\mathbf{B}_t - \frac{1}{2} \int_0^T \mathbf{h}_t \mathbf{h}_t' dt\right)$$

- is the Radon-Nikodym derivative.
- The discrepancy between Q and P can be measured by the *relative entropy* between Q^* and P^* , defined as

$$\begin{aligned}\mathcal{R}(Q) &= \delta \int_0^\infty \exp(-\delta t) E_Q(\log(\xi_t)) dt \\ &= \delta \int_0^\infty \exp(-\delta t) E_Q\left(\int_0^t \mathbf{h}_\tau d\mathbf{B}_\tau - \frac{1}{2} \int_0^t \mathbf{h}_\tau \mathbf{h}'_\tau d\tau\right) dt \\ &= \int_0^\infty \exp(-\delta \tau) E_Q\left(\frac{1}{2} \mathbf{h}_\tau \mathbf{h}'_\tau\right) d\tau\end{aligned}$$

- This measure has the nice property to measure a relative distance between two distributions.
- It is zero if and only if $Q = P$, and it is otherwise positive.
- Notice that $Q = P$ coincide here when $\mathbf{h}_\tau = 0$ (i.e. $\mathbf{B} = \widehat{\mathbf{B}}$).
- The *multiplier robust control problem* is then formulated as

$$\sup_{c, \varphi} \inf_Q \left\{ E_Q \left[\int_0^T e^{-\delta t} u(c_t) dt \right] + \theta \mathcal{R}(Q) \right\}$$

- subject to the state equation (budget constraint) (26).
- Here θ is a penalty imposed on the discrepancy between Q and P . It is just one more parameter to be chosen to “calibrate” the model.
- Notice that a high θ implies a choice of Q that is close to P because we are taking the “inf” with respect to Q .
- In other words, we are giving most of the weight to the relative entropy and little to the utility function.

- If $\theta = 0$, we consider all the possible Q 's.
- It turns out that the following representation also holds:
- The *multiplier robust control problem* can be formulated as

$$\sup_{c, \varphi} \inf_{\mathbf{h}} \left\{ \widehat{E} \left[\int_0^T e^{-\delta t} \left(u(c_t) + \frac{\theta}{2} \mathbf{h}_t \mathbf{h}_t' \right) dt \right] \right\}$$

- subject to

$$dW_t = \mu_W(W_t, \mathbf{X}_t, \varphi_t, c_t) dt + \sigma_W(W_t, \mathbf{X}_t, \varphi_t, c_t) (d\widehat{\mathbf{B}}_t + \mathbf{h}_t dt) \quad (31)$$

- where \widehat{E} denote the expectation computed by integrating the Brownian motion $\widehat{\mathbf{B}}_t$.
- A necessary condition turns out to be the Bellman Isaacs condition:
- To formulate it, it is convenient to stack all the state variables. Define $\mathbf{Y}_t = (W_t, \mathbf{X}_t)'$, so that we have

$$d\mathbf{Y}_t = \boldsymbol{\mu}_Y(\mathbf{Y}_t, t, \varphi_t, c_t) dt + \boldsymbol{\sigma}_Y(\mathbf{Y}_t, t, \varphi_t, c_t) (d\widehat{\mathbf{B}}_t + \mathbf{h}_t dt)$$

- The Bellman Isaacs condition is then the following:
- There exists a value function $J(Y)$ such that

$$\begin{aligned} \delta J &= \max_{c, \varphi} \min_{\mathbf{h}} \left\{ u(c) + \frac{\theta}{2} \mathbf{h} \mathbf{h}' + (\boldsymbol{\mu}_Y + \boldsymbol{\sigma}_Y \mathbf{h}')' \mathbf{J}_Y \right. \\ &\quad \left. + \frac{1}{2} \text{tr}(\boldsymbol{\sigma}_Y' \mathbf{J}_{YY} \boldsymbol{\sigma}_Y) \right\} \\ &= \min_{\mathbf{h}} \max_{c, \varphi} \left\{ u(c) + \frac{\theta}{2} \mathbf{h} \mathbf{h}' + (\boldsymbol{\mu}_Y + \boldsymbol{\sigma}_Y \mathbf{h}')' \mathbf{J}_Y + \frac{1}{2} \text{tr}(\boldsymbol{\sigma}_Y' \mathbf{J}_{YY} \boldsymbol{\sigma}_Y) \right\} \end{aligned} \quad (32)$$

- Solving for the minimum \mathbf{h} , one obtains

$$\mathbf{h}' = -\frac{1}{\theta} \boldsymbol{\sigma}'_Y \mathbf{J}_Y$$

- Notice that then

$$\begin{aligned} \frac{\theta}{2} \mathbf{h} \mathbf{h}' &= \frac{1}{2\theta} \mathbf{J}'_Y \boldsymbol{\sigma}_Y \boldsymbol{\sigma}'_Y \mathbf{J}_Y \\ \boldsymbol{\sigma}_Y \mathbf{h}' &= -\frac{1}{\theta} \boldsymbol{\sigma}_Y \boldsymbol{\sigma}'_Y \mathbf{J}_Y \end{aligned}$$

- Substitute into (32) to find

$$\delta J = \max_{c, \boldsymbol{\varphi}} \left\{ u(c) - \frac{1}{2\theta} \mathbf{J}'_Y \boldsymbol{\sigma}_Y \boldsymbol{\sigma}'_Y \mathbf{J}_Y + \mu'_Y \mathbf{J}_Y + \frac{1}{2} \text{tr}(\boldsymbol{\sigma}'_Y \mathbf{J}_{YY} \boldsymbol{\sigma}_Y) \right\} \quad (33)$$

- Since consumption is only affecting the “wealth” state variable, we obtain that the usual first order conditions for consumption yields

$$u_c = J_W$$

- Instead, the FOC for optimal portfolio weights imply

$$\begin{aligned} 0 &= (\boldsymbol{\mu}_S - r \mathbf{1}_d) J_W + J_{WW} W_t \boldsymbol{\sigma}_S \boldsymbol{\sigma}'_S \boldsymbol{\varphi}'_t + \boldsymbol{\sigma}_S \mathbf{a}' \mathbf{J}_{W\mathbf{x}} \\ &\quad - \frac{1}{2\theta} J_W^2 W \boldsymbol{\sigma}_S \boldsymbol{\sigma}'_S \boldsymbol{\varphi}'_t - \frac{1}{\theta} J_W \boldsymbol{\sigma}_S \mathbf{a}' \mathbf{J}_{\mathbf{x}} \end{aligned}$$

- The latter yields

$$\begin{aligned} \varphi_t = & \frac{-J_W}{W_t \left(J_{WW} - \frac{1}{\theta} J_W^2 \right)} \left(\boldsymbol{\sigma}_{S,t} \boldsymbol{\sigma}'_{S,t} \right)^{-1} \left(\boldsymbol{\mu}_{S,t} - r_t \mathbf{1}_d \right) \\ & + \frac{-1}{W_t \left(J_{WW} - \frac{1}{\theta} J_W^2 \right)} \left(\boldsymbol{\sigma}_{S,t} \boldsymbol{\sigma}'_{S,t} \right)^{-1} \boldsymbol{\sigma}_{S,t} \mathbf{a}' \mathbf{J}_{W\mathbf{x}} \\ & + \frac{\frac{1}{\theta} J_W}{W_t \left(J_{WW} - \frac{1}{\theta} J_W^2 \right)} \left(\boldsymbol{\sigma}_{S,t} \boldsymbol{\sigma}'_{S,t} \right)^{-1} \boldsymbol{\sigma}_{S,t} \mathbf{a}' \mathbf{J}_{\mathbf{x}} \end{aligned}$$

- The portfolio rule has then three component: The first is the standard myopic demand. Notice that the denominator is adjusted for robustness, implying a lower investment in the stocks (because $J_W^2 \frac{1}{\theta} > 0$).
- The second is the standard Merton's hedging demand (recall that $\mathbf{a}(\mathbf{X}, \mathbf{t})$ is the diffusion term of the state variables \mathbf{X}).
- The third is also a hedging demand, which arises because of robustness preferences.

5.2 Homothetic Preferences

- Maenhout (2004) introduces a simple rescaling of $1/\theta$ to obtain back homotheticity of the preferences under robust control.
- Let us focus to the simple case where there are no state variables \mathbf{X} , so that stock returns have constant parameters

$$d\mathbf{S}_t = \mathbf{I}_S \boldsymbol{\mu}_S dt + \mathbf{I}_S \boldsymbol{\sigma}_S d\mathbf{B}_t \quad (34)$$

- and suppose that

$$u(c_t) = e^{-\phi t} \frac{c_t^{1-\gamma}}{1-\gamma}$$

- We then have that the FOC of the Bellman-Isaac condition are

$$\mathbf{h}' = -\frac{1}{\theta} \boldsymbol{\sigma}'_W J_W$$

- Hence

$$\delta J = \max_{c, \varphi} \left\{ u(c) - \frac{1}{2\theta} J_W^2 \boldsymbol{\sigma}_W \boldsymbol{\sigma}'_W + \mu_W J_W + \frac{1}{2} J_{WW} \boldsymbol{\sigma}_W \boldsymbol{\sigma}'_W \right\} \quad (35)$$

- Since consumption is only affecting the “wealth” state variable, we obtain that the usual first order conditions for consumption yields

$$u_c = J_W$$

- The latter yields

$$\boldsymbol{\varphi}_t = \frac{-J_W}{W_t \left(J_{WW} - \frac{1}{\theta} J_W^2 \right)} (\boldsymbol{\sigma}_S \boldsymbol{\sigma}'_S)^{-1} (\boldsymbol{\mu}_S - r \mathbf{1}_d)$$

- It is easy to see that with a constant $1/\theta$, preferences would not be homothetic (or scale invariant).

- In fact, if they were, we could write $J(W, t) = k(t) W^{1-\gamma}$ (where $k < 0$ if $\gamma > 1$), so that $J_W^2 = k(t)^2 (1-\gamma)^2 W^{-2\gamma}$ while $J_{WW} = -k(t) \gamma (1-\gamma) W^{-\gamma-1}$.
- This implies immediately that

$$\varphi_t = \frac{1}{\left(\gamma + \frac{1}{\theta} k(t) (1-\gamma) W_t^{1-\gamma}\right)} (\boldsymbol{\sigma}_S \boldsymbol{\sigma}'_S)^{-1} (\boldsymbol{\mu}_S - r \mathbf{1}_d)$$

- Clearly, the portfolio weights are not “scale invariant” because they depend on W_t (unless $\gamma = 1$).
- Maenhout (2004) proposes to scale the penalty parameter θ by the value function J itself, in a way to make the model again scale independent.
- Specifically, he assumes

$$\theta = \theta(J) = \theta^* (1-\gamma) J(W, t)$$

- Hence, the optimal level of perturbation \mathbf{h} becomes

$$\mathbf{h}' = -\frac{1}{\theta^* (1-\gamma) J(W, t)} \boldsymbol{\sigma}'_W J_W$$

- The rescaling of the optimal perturbation \mathbf{h} by the value function J makes sure that as W increases, \mathbf{h} remains a first order effect.

- This can also be seen when we substitute \mathbf{h} back into the Bellman-Isaac equation

$$\delta J = \max_{c, \varphi} \left\{ u(c) + \frac{1}{2} \boldsymbol{\sigma}_W \boldsymbol{\sigma}'_W \left(J_{WW} - \frac{J_W^2}{\theta^* (1 - \gamma) J} \right) + \mu_W J_W \right\} \quad (36)$$

- Maenhout (1999) then obtains the following results (this is just the obvious generalization to multiple assets):

- The value function is given by

$$J(W, t) = \left[\frac{1 - e^{-a(T-t)}}{a} \right] \frac{W^{1-\gamma}}{1-\gamma}$$

- where

$$a = \frac{1}{\gamma} \left[\phi - (1 - \gamma)r - \frac{1 - \gamma}{2(\gamma + \theta)} (\boldsymbol{\mu}_S - r\mathbf{1}_d)' (\boldsymbol{\sigma}_S \boldsymbol{\sigma}'_S)^{-1} (\boldsymbol{\mu}_S - r\mathbf{1}_d) \right]$$

- and

$$c_t = \frac{a}{1 - e^{-a(T-t)}} W_t$$

$$\varphi_t = \frac{1}{\gamma + 1/\theta^*} (\boldsymbol{\sigma}_S \boldsymbol{\sigma}'_S)^{-1} (\boldsymbol{\mu}_S - r\mathbf{1}_d)$$

5.3 Asset Pricing Implications

- Consider an economy with aggregate dividends

$$\frac{dD_t}{D_t} = \mu_D dt + \sigma_D dB_t$$

- where σ_e is a constant and B is a BM.
- The stock return on the market is given by

$$dR = \frac{dS_t + D_t dt}{S_t}$$

- and this will follow the Ito process

$$dR = \mu_R dt + \sigma_R dB_t$$

- where μ_R and σ_R have to be determined in equilibrium.
- We can conjecture they are constant.
- If they are, then applying the results of the previous section and imposing the market clearing condition $\varphi_t = \mathbf{1}_d$, we obtain

$$\mu_R - r = \left(\gamma + \frac{1}{\theta^*} \right) \sigma_R^2$$

- Notice in addition that (assuming infinite horizon)

$$c_t = D_t = aW_t$$

- Since there is no other wealth than the financial wealth in the economy, we must then have that

$$W_t = S_t$$

- This implies

$$S_t = aD_t = ac_t$$

- Hence

$$\sigma_R = \sigma_c$$

- Therefore, we can also write

$$\mu_R - r = \left(\gamma + \frac{1}{\theta^*} \right) \text{cov} (dR_t, dc_t/c_t)$$

- We can increase the equity risk premium by decreasing the (relative) penalty θ^* on the possible distributions.
- Differently from increases in risk aversion, decreasing θ^* also decreases the risk-free rate, avoiding the “risk-free rate puzzle”

$$r = \phi + \gamma\mu_D - \frac{1}{2} (1 + \gamma) \left(\gamma + \frac{1}{\theta^*} \right) \sigma_D^2$$

- Hence, a decrease in θ^* indeed decreases the risk-free rate through the precautionary saving motives.
- In other words, preferences for robustness introduce one more parameter which is delinked from the intertemporal elasticity of substitution $1/\gamma$.
- As with EZ preferences, we can then increase the “aversion to model uncertainty” without affecting the intertemporal substitution, thereby obtaining a high equity risk premium and a low interest rate.

- Finally, notice that indeed r , μ_R and σ_R are constants, so that the initial conjecture is verified.

5.4 Calibrating Pessimism

- Recall that “pessimism” stems from the perturbation of the Brownian motions

$$dB_t = d\widehat{B}_t + hdt$$

- That is, we must substitute dB_t with $d\widehat{B}_t + hdt$ in

$$dR_t = \mu_R dt + \sigma_R dB_t$$

- to obtain

$$dR_t = (\mu_R + \sigma_R h) dt + \sigma_R d\widehat{B}_t$$

- The optimal robustness parameter was found to be

$$\begin{aligned} h &= -\frac{1}{\theta^* (1 - \gamma) J} \sigma_W J_W \\ &= -\frac{1}{\theta^* (1 - \gamma) J} \sigma_W \frac{1}{a} W^{-\gamma} \\ &= -\frac{1}{\theta^*} \sigma_c \end{aligned}$$

- Using the results in the previous subsection, we have

$$\mu_R^* = \mu_R + \sigma_R h = \left(\gamma + \frac{1}{\theta^*} \right) \sigma_c^2 - \frac{1}{\theta^*} \sigma_D^2 = \gamma \sigma_c^2$$

- μ_R^* is the worst case equity premium (that is, adjusted for the “optimal” perturbation h).
- In a partial equilibrium model, this equals the actual one μ_R plus a (negative) term $\sigma_R h$ that make agents decide to decrease the demand for assets.
- In general equilibrium, this worst case scenario implies $\mu_R^* = \gamma \sigma_D^2$, that is, the value under standard expected utility.
- Notice that this is independent of θ^* . Intuitively, in a general equilibrium framework, a decrease in θ^* increases both the equity premium μ_R and the aversion to misspecification $\sigma_R h$ (which is negative) by the same amount.
- This implies that robust preferences cannot explain the equity risk premium puzzle! The pessimistic μ_R^* is simply too pessimistic, given by $\gamma \sigma_c^2$ (which is small if γ is kept small to maintain the risk free rate small).

5.5 The Link with SDUs

- One final comment is the observational equivalence between robustness and SDUs.
- Consider the simple case with only wealth as state variable.
- We can rewrite (35)

$$0 = \max_{c, \varphi} \left\{ u(c) - \delta J + \mathcal{D}^{(c, \varphi)} J + \frac{1}{\theta} J_W^2 \sigma_W \sigma'_W \right\} \quad (37)$$

- where

$$\mathcal{D}^{(c, \varphi)} J = \mu_W J_W + \frac{1}{2} J_{WW} \sigma_W \sigma'_W \quad (38)$$

- (37) has the same form of the Hamilton - Jacobi - Bellman equation of the stochastic differential utility, with aggregator $f(c, J) = u(c) - \delta J$ and variance multiplier $A(J) = 1/\theta$, that is, the parameter for robustness.
- In fact, if $J(W)$ is the utility function (as a function of the unique state variable W : Recall that also c depends on W), notice that its volatility $\sigma_J = J_W \sigma_W$.
- The standard time-separable utility is then given by the case $\theta = \infty$, which imply that the last term cancel out.
- This result implies an observational equivalence between robustness and recursive utility.

6 Recent Applications

- The approach of ambiguity aversion, or robust control theory, have found a number of applications in finance in recent times.
- For instance, in a robust control setting as in Anderson, Hansen and Sargent (1998), Liu, Pan and Wang (2005) study uncertainty on rare events only (that is, the set of prior probabilities only contain the probability of a jump in the underlying endowment process), and apply the theory to explain options premia, along with the standard result on returns equity premium.
- Similarly, Routledge and Zin (2004) study the implications of rare events for market liquidity, and show that uncer-

tainty aversion may lead agents not to trade, after big market events.

- Finally, Uppal and Wang (2003) extend the modeling device for robust control, and specifically the re-scaling technique of Maenhout (2004), to study the case of different aversions to uncertainty across a set of assets. For instance, for some assets one may feel more “certain” about (e.g. there is more information) while for others one may feel less “certain.” They use the model to “explain” under-diversification: even a limited amount of aversion to uncertainty on some stocks make the investors over-invest in those for which there is less uncertainty aversion. Perhaps this model may have interesting implications for the Cross-Section of stock returns.

7 References

- Anderson, Hansen and Sargent (2000) “Robustness, Detection and the Price of Risk” WP, University of Chicago
- Dow and Werlang (1992) “Uncertainty Aversion, Risk Aversion and the Optimal Choice of Portfolio), *Econometrica*.
- Cagetti, Hansen, Sargent and Williams (2002) “Robustness and Pricing with Uncertain Growth” *Review of Financial Studies*, University of Chicago;
- Chen and Epstein (2002) “Ambiguity, Risk and Asset Returns in Continuous Time,” *Econometrica*
- Epstein and Wang (1994) “Intertemporal Asset Pricing under Knightian Uncertainty” *Econometrica*

- Hansen, Sargent, Turnmuhambetova and Williams (2001) “Robustness and Uncertainty Aversion” WP, University of Chicago;
- Hansen and Sargent (2001) “Robust Control and Model Uncertainty” WP, University of Chicago;
- Liu, Pan and Wang (2005) “An Equilibrium Model of Rare Event Premia”, RFS.
- Maenhout (2004) “Robust Portfolio Rules and Asset Pricing” RFS
- Routledge and Zin (2004) “Model Uncertainty and Liquidity”, WP, NBER
- Uppal and Wang (2003) “Model misspecification and underdiversification”, JF