Lecture 4: Non-expected utility in asset pricing

I The macroeconomic announcement premium and NonEU

The macro announcement premium AB (2018): empirical evidence

A-SDF: a two-period example AB (2018) Supplemental material

Theorem of GRS In this section, we consider a very general class of intertemporal preferences that has the recursive representation

$$V_t = \mathcal{I}\left[u\left(C_t\right) + \beta V_{t+1}\right], \tag{1}$$

where $\beta \in (0,1)$ and $\mathcal{I}$ is the certainty equivalence functional that maps the next-period utility (which is a random variable) into its certainty equivalent (which is a real number).

Definition 1 *(Generalized Risk Sensitivity)*

An intertemporal preference of the form (1) is said to satisfy (strictly) generalized risk sensitivity, if the certainty equivalence functional $\mathcal{I}$ is (strictly) monotone with respect to second-order stochastic dominance.

Here we provide a proof for Theorem 2 of the paper in a two-period model by assuming i) a finite state space, ii) fully revealing announcements, and iii) equal probability of each state. The proof in the paper shows that the conclusion of the theorem holds in a fully dynamic model without assuming fully revealing announcements. In addition, the assumption of a finite space with equal probability can be replaced by a continuum. Because the proof of Theorem 2 under the assumption of finite state space is relatively simple and does not require functional analysis in infinite dimensional spaces, we present such a proof in this note. As in the paper, we make the following assumptions on $\mathcal{I}$:

Assumption 1: $\mathcal{I}$ is continuously differentiable with strictly positive partial derivatives.
Assumption 2: $\mathcal{I}[k] = k$ whenever k is a constant.

As we show in equation (12) and (13) on page 9 of the paper,

$$P^- = E\left[m^*(s)\, P^+(s)\right], \tag{2}$$

where the A-SDF, $m^*(s)$ is given by:

$$m^*(s) = \frac{1}{\pi(s)} \frac{\frac{\partial}{\partial V_s}\mathcal{I}[V]}{\sum_{n=1}^N \frac{\partial}{\partial V_n}\mathcal{I}[V]}. \tag{3}$$

In the above equation, $\frac{\partial}{\partial V_s}\mathcal{I}[V]$ denotes the partial derivative of $\mathcal{I}[V]$ with respect to its *sth* element. Equation (2) implies that the announcement premium is positive (negative) if

$$E\left[m^*(s)P^+(s)\right] \leq (\geq) E\left[P^+(s)\right].$$

We first show that Condition 1 in the paper is equivalent to the "negative comonotonicity" of the partial derivatives of $\mathcal{I}[V]$:

Lemma 1 *The following two conditions are equivalent:*

1. *The announcement premium is non-negative for all payoffs that are comonotone with V.*[1]

2. *For any $V \in \Psi^N$,*
$$\left(\frac{\partial}{\partial V_s}\mathcal{I}[V] - \frac{\partial}{\partial V_{s'}}\mathcal{I}[V]\right)(V_s - V_{s'}) \leq 0. \tag{4}$$

Proof. *First, we assume that 1) is true and prove 2) by contradiction. Suppose there exist V and s, s' such that $V_s > V_{s'}$ and $\frac{\partial}{\partial V_s}\mathcal{I}[V] > \frac{\partial}{\partial V_{s'}}\mathcal{I}[V]$. Consider the following payoff:*

$$X(n) = V_n \quad for \quad n = s, s'; \qquad X(n) = 0 \qquad otherwise.$$

Clearly, X is comonotone with V, and therefore positively correlated with $m^*(s)$ defined in (3). Therefore,

$$P^- = E[m^*(s)X(s)] > E[m^*(s)]E[X(s)] = E[X(s)],$$

contradicting a non-negative announcement premium.

Next, we assume that 2) is true and prove 1). Take any X that is comonotone with V, then

$$P^- = E[m^*(s)X(s)] \leq E[m^*(s)]E[X(s)] = E[X(s)]$$

because $m^*(s)$ and $X(s)$ are negatively correlated. ■

[1] Recall that a payoff X is comonotone with V if $\forall s$ and s' such that $X(s) \cdot X(s') \neq 0$, $[X(s) - X(s')][V(s) - V(s')] \geq 0$.

Lemma 1 establishes the equivalence between non-negative announcement premium (for payoffs that are comonotone with continuation utility) and inequality (4). Inequality (4) is known to be a characterization of Schur concave functions, which is equivalent to monotone with respect to second order stochastic dominance for functions defined on finite probability spaces with equal probabilities. We summarize the equivalence results in the following lemma and refer the readers to Marshal and Okin or Muller and Stoyan for reference of such results.

Lemma 2 *For any $\mathcal{I}$ that satisfies Assumption 1, the following two statements are equivalent:*

1. *$\mathcal{I}[V]$ is non-decreasing in second order stochastic dominance if and only if for any $V \in \Psi^N$, $\left(\frac{\partial}{\partial V_s}\mathcal{I}[V] - \frac{\partial}{\partial V_{s'}}\mathcal{I}[V]\right)(V_s - V_{s'}) \leq 0$.*
2. *$\mathcal{I}[V]$ is strictly increasing in second order stochastic dominance if and only if any $V \in \Psi^N$, $\left(\frac{\partial}{\partial V_s}\mathcal{I}[V] - \frac{\partial}{\partial V_{s'}}\mathcal{I}[V]\right)(V_s - V_{s'}) \leq 0$, and strict inequality holds whenever $V_s \neq V_{s'}$.*
3. *$\mathcal{I}[V]$ is non-increasing in second order stochastic dominance if and only if any $V \in \Psi^N$, $\left(\frac{\partial}{\partial V_s}\mathcal{I}[V] - \frac{\partial}{\partial V_{s'}}\mathcal{I}[V]\right)(V_s - V_{s'}) \geq 0$.*

With the above we are ready to prove Theorem 2 in the paper. The first part of Theorem 2 is

1. *The announcement premium is zero for all assets if and only if $\mathcal{I}$ is expected utility.*

   Proof. If $\mathcal{I}$ is the expectation operator, that is, $\mathcal{I}[V] = \sum_{s=1}^{N} \pi(s) V(s)$, then by (3), $m^*(s) = 1$, and the announcement premium must be zero for all assets. Conversely, if the announcement premium is zero for all assets, we must have $m^*(s) = m^*(s')$ for all s, s', otherwise we can construct an asset with nonzero payoff in state s and s' that requires a non-trivial announcement premium. This implies that $\left(\frac{\partial}{\partial V_s}\mathcal{I}[V] - \frac{\partial}{\partial V_{s'}}\mathcal{I}[V]\right)(V_s - V_{s'}) = 0$ for all s,s'. For any $V \in \Psi^N$, note that $E[V] \geq_{SSD} V$, by the above lemma, we must have

   $$\mathcal{I}[V] = \mathcal{I}[E[V]] = E[V],$$

   where the last equality uses Assumption 2. ■

   The second part of Theorem 2 is a direct consequence of Lemma 1 and Lemma 2:

2. *The announcement premium is non-negative for all assets with payoffs comonotone with V if and only $\mathcal{I}$ is non-decreasing with respect to second order stochastic dominance.*

   From the above discussion, it is clear that a stronger version of the above result is also true, that is,

3. *The announcement premium is strictly positive for all assets with payoffs strongly comonotone with V if and only $\mathcal{I}$ is strictly increasing with respect to second order stochastic dominance.*

II Recursive utility

A From Bellman equation to recursive utility

- The addatively separable expected utility is recursive:

$$V(t) = u(C_t) + \beta E[V(t+1)].$$

- Bellman equation goes from the infinite sum to the recursive representation. We can start from the latter! In general, a recursive preference can be described by a pair of aggregators, W and m. The function $m(\cdot)$ is called the certainty equivalence functional. In general, $m(\cdot)$ can be constructed using an strictly increasing and concave function h:

$$m(X) = h^{-1}\{E[h(X)]\}.$$

  The KP preference features CRRA certainty equivalence functional, i.e. $h(X) = \frac{1}{1-\gamma}X^{1-\gamma}$. In this case,

$$m(X) = \left\{E\left[X^{1-\gamma}\right]\right\}^{\frac{1}{1-\gamma}}.$$

  The function W is the intertemporal aggregator. KP uses the CES intertemporal aggregator:

$$W(c, m) = \left\{(1-\beta)c^{1-\frac{1}{\psi}} + \beta m^{\frac{1}{1-\psi}}\right\}^{\frac{1}{1-1/\psi}}.$$

B Preference for early resolution of uncertainty

The Kreps-Porteus example Consider two economies with identical rep agents and identical consumption processes. The only difference is that in Economy L, the uncertainty of consumption in period 1 is resolved only in period 1, while the uncertainty in Economy E is resolved in period 0.

- Economy L: Late resolution of uncertainty

	Period 0		Period 1
π_H	C_0	Info	C_H
π_L		Arrive	C_L

(5)

- Economy E: Early resolution of uncertainty

		Period 0	Period 1
π_H	Info	C_0	C_H
π_L	Arrive	C_0	C_L

(6)

- In period -1, nature decides the consumption path, the probability is given by π_i, for $i = H, L$. In both economies, there is no uncertainty with respect to date 0 consumption. The only uncertainty is about date one consumption.

- In Economy L, the consumption path is not revealed to the agent until the end of peiod 1. In Economy E, the consumption path is revealed to the agent immediately. Krep and Porteus use a similar example to illustrate preference for early resolution.

- In the next sections, I briefly review the KP definition of early (late) resolution. I then derive the SDF of both economies. Finally, I prove that the in economy E, news announcement requires a positive premium if and only if agents prefers early resolution.

Preference for Early Resolution Under the above notation, the date-0 utility of the rep agent of economy L can be calculated as:

$$U_L = W\left\{C_0, h^{-1}\left(E\left[h\left(C_1\right)\right]\right)\right\},$$

and the date-0 utility of the agent in economy E can be calculated as:

$$U_E = h^{-1}\left\{E\left[h \circ W\left\{C_0, C_1\right\}\right]\right\}.$$

Because h is strictly increasing, $U_L < U_E$ if and only if $h(U_L) < h(U_E)$, which is

$$h \circ W\left\{C_0, h^{-1}\left(E\left[h\left(C_1\right)\right]\right)\right\} < E\left[h \circ W\left\{C_0, C_1\right\}\right]. \tag{7}$$

Denote

$$x = h\left(C_1\right),$$

and

$$f(x) = h \circ W\left\{C_0, h^{-1}(x)\right\}. \tag{8}$$

Then the left-hand of (7) is

$$f\left(E[x]\right),$$

and the right-hand side of (7) is;

$$E\left[f(x)\right].$$

This implies that $U_L < U_E$ if and only if the f function defined in (8) is strictly convex. We make the following observation.

Claim 1 *(Preference for Early Resolution)*

The recursive preference (W, m) exhibits preference for early (late) resolution of uncertainty if and only if the f function defined in (8) is strictly convex (concave).

Under the KP preference, this is equivalent to $\gamma > \frac{1}{\psi}$ (or $\gamma < \frac{1}{\psi}$).

Note that under the KP preference,

$$f(x) = \frac{1}{1-\gamma}\left\{(1-\beta) C_0^{1-\frac{1}{\psi}} + \beta\left[(1-\gamma) x\right]^{\frac{1-1/\psi}{1-\gamma}}\right\}^{\frac{1-\gamma}{1-1/\psi}}.$$

One can show that

$$f''(x) > 0 \text{ if and only if } \gamma > \frac{1}{\psi}.$$

C The SDF for recursive utility

$$V(W_t) = \max_{C_t, \alpha_0}\left\{(1-\beta) C_t^{1-\frac{1}{\psi}} + \beta E\left[V\left(W_{t+1}\right)^{1-\gamma}\right]^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\right\}^{\frac{1}{1-\frac{1}{\psi}}}$$

$$\text{s.t.} \quad C_t + \alpha = e_t + W_t$$

$$W_{t+1} = \alpha R_{t+1} + e_{t+1}$$

where α is the share of the portfolio which agent buys at time t, R_{t+1} is the return for the portfolio at time $t+1$ and e_t, e_{t+1} is the endowment at time t and $t+1$ respectively.

Let us put the budget constraint into the utility function, we can get

$$V(W_t) = \left\{ (1-\beta)\left(e_t + W_t - \alpha\right)^{1-\frac{1}{\psi}} + \beta E\left[V\left(\alpha R_{t+1} + e_{t+1}\right)^{1-\gamma}\right]^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right\}^{\frac{1}{1-\frac{1}{\psi}}}$$

The first order condition of α :

$$(1-\beta)C_t^{-\frac{1}{\psi}} = \beta m^{-\left(\frac{1}{\psi}-\gamma\right)} E\left[V\left(W_{t+1}\right)^{-\gamma} V_W\left(W_{t+1}\right) R_{t+1}\right] \tag{9}$$

where $m = E\left[V_{t+1}^{1-\gamma}\right]^{\frac{1}{1-\gamma}}$.

The envelop condition:

$$V_W = (1-\beta)\, V\left(W_t\right)^{\frac{1}{\psi}} C_t^{-\frac{1}{\psi}} \tag{10}$$

Combining equation (9) and equation (10), we can get

$$\begin{aligned}(1-\beta)C_t^{-\frac{1}{\psi}} &= \beta m^{-\left(\frac{1}{\psi}-\gamma\right)} E\left[V\left(W_{t+1}\right)^{-\gamma}(1-\beta)\, V\left(W_{t+1}\right)^{\frac{1}{\psi}} C_{t+1}^{-\frac{1}{\psi}} R_{t+1}\right] \\ 1 &= E\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{1}{\psi}}\left(\frac{V\left(W_{t+1}\right)}{m}\right)^{\frac{1}{\psi}-\gamma} R_{t+1}\right]\end{aligned}$$

Therefore, the pricing kernel $M_{t+1} = \beta\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{1}{\psi}}\left(\frac{V_{t+1}}{E\left[V_{t+1}^{1-\gamma}\right]^{\frac{1}{1-\gamma}}}\right)^{\frac{1}{\psi}-\gamma}$

III Multiple-prior (Maxmin)

IV Robust control

A Entropy and relative entropy

Entropy Entropy is a concept that is original used in thermodynamics and information theory. It has found some interesting applications in economics (related to information) as well. I begin by introducing some terminologies and definitions.

Let Ω be a finite set and $p(\omega)$ be a probability mass function on Ω. The entropy of p is defined by:

$$H(p) = -\sum p(\omega) \ln p(\omega).$$

Several comments:

- If $p(\omega) = 0$, we adopt the convention $p(\omega)\ln p(\omega) = 0$. The rational is $\lim_{p\to 0} p \ln p = 0$.

- Because $p(\omega) \leq 1$, $\ln p(\omega) < 0$ for all ω. It follows immediately that $H(p) \geq 0$. Note that if p is a continuous density, this may no longer hold. That is, $H(p) = -\int p(\varepsilon)\ln p(\varepsilon)\, d\omega < 0$ is possible. A trivial example is a uniform distribution on $[0, a]$, for $a \in (0, 1)$.

- We can similiarly define the entropy of a random variable. Let $X : \Omega \to R$, then

$$H(X) = -\sum_x P(X = x) \ln P(X = x).$$

  An alternative way to think about the entropy of a random variable is to think of p as the probability mass function of X and view $p(X)$ as a random variable. We can simply write

$$H(X) = -E[\ln p(X)].$$

Relative entropy Let p and q be two probability measure on Ω, the relative entropy of p with respective q is defined as:

$$D(p\| q) = \sum p(\omega) \ln \frac{p(\omega)}{q(\omega)}.$$

Several remarks:

- Note that

$$D(p\| q) = \sum p(\omega) \ln \frac{p(\omega)}{q(\omega)} \neq D(q\| p) = \sum q(\omega) \ln \frac{q(\omega)}{p(\omega)}$$

- It is easy to see that relative entropy must be non-negative:

$$\begin{aligned} D(p\| q) &= \sum p(\omega) \ln \frac{p(\omega)}{q(\omega)} = -\sum p(\omega) \ln \frac{q(\omega)}{p(\omega)} \\ &\geq -\ln \sum p(\omega) \times \frac{q(\omega)}{p(\omega)} \quad (Jensen's\ inequality) \\ &= 0. \end{aligned}$$

- An alternative way to say this is

$$H(p) = -\sum p(\omega)\ln p(\omega) \leq \sum p(\omega)\ln q(\omega),$$

  for all q, and equality achieves only if $p = q$.

- Again, if X and Y are two random variables, then

$$D(X\| Y) = E\left[\ln\frac{p(X)}{p(Y)}\right] = \sum_{x,y} p(x,y)\ln\frac{p(x)}{p(y)}.$$

Entropy of a strictly positive random variable For a strictly positive random variable X, sometimes the entropy of X is defined as

$$\mathcal{L}(X) = \ln E[X] - E[\ln X].$$

This can be motivated as follows. To provide a complete decription, let's say X is an rv defined on a probability space $(\Omega, \mathcal{F}, P)$. Because X is strictly positive, $\frac{X}{E[X]}$ is a density. That is, we can define a probabiliity measure Q out of $\frac{X}{E[X]}$ as

$$Q(A) = \int \frac{X}{E[X]} I_A dP, \quad A \in \mathcal{F}$$

Now let's compute the relative entropy $D(P\| Q)$:

$$D(P\| Q) = \int \ln\left[\frac{X}{E[X]}\right] dP = \ln E[X] - E[\ln X].$$

Note that entropy can be viewed as a measure of how dispersed X is. This is particularly useful when higher moments of X does not exist.

An example is a log normal distribution. Let $\ln X \sim N(\mu, \sigma^2)$. Then $E[X] = e^{\mu + \frac{1}{2}\sigma^2}$, and $\mathcal{L}(X) = \ln E[X] - E[\ln X] = \frac{1}{2}\sigma^2$. Recall that in the international context

$$\begin{aligned} E[\ln \Delta e_{t+1}] &= E\left[\ln m^*_{t+1}\right] - E[\ln m_{t+1}] \\ &= \mathcal{L}(m_{t+1}) - \mathcal{L}(m^*_{t+1}) + r_t - r^*_t. \end{aligned}$$

Let $\ln m_{t+1} \sim N(-\mu_t, \sigma_t^2)$, and $\ln m^*_{t+1} \sim N(-\mu^*_t, \sigma_t^{*2})$, we have:

$$r_t = \ln\frac{1}{E[m_{t+1}]} = \mu_t - \frac{1}{2}\sigma_t^2; \quad r^*_t = \ln\frac{1}{E\left[m^*_{t+1}\right]} = \mu^*_t - \frac{1}{2}\sigma_t^{*2}.$$

We have:

$$\begin{aligned} E\left[\ln \Delta e_{t+1}\right] &= E\left[\ln m^*_{t+1}\right] - E\left[\ln m_{t+1}\right] = \mu_t - \mu^*_t; \\ \mathcal{L}(m_{t+1}) - \mathcal{L}(m^*_{t+1}) &= \frac{1}{2}\left(\sigma^2_t - \sigma^{*2}_t\right); \\ r_t - r^*_t &= \left(\mu_t - \mu^*_t\right) - \frac{1}{2}\left(\sigma^2_t - \sigma^{*2}_t\right). \end{aligned}$$

Therefore, the exchange no arbitrage condition can be written as:

$$\left[r_t - r^*_t\right] + \frac{1}{2}\left(\sigma^2_t - \sigma^{*2}_t\right) = \underbrace{\frac{1}{2}\left(\sigma^2_t - \sigma^{*2}_t\right)} + \underbrace{\left[r_t - r^*_t\right]}$$

Hence,

$$Cov\left[E_t\left[\ln \Delta e_{t+1}\right], \left[r_t - r^*_t\right]\right] = Var\left[r_t - r^*_t\right] + \frac{1}{2} Cov\left[\sigma^2_t - \sigma^{*2}_t, r_t - r^*_t\right]$$

A.1 The Robust Control Model

This section is a brief summary of the link/comparison between Hansen Sargent and Gilboa and Schmeidler.

One can think of robust control as a special case of the Gilboa-Schmeidler multiple prior utility model. An expected utility agent maximizes:

$$\sum_{i=1}^{N} \pi\left(i\right) u\left(C_i\right),$$

where $\pi = [\pi_1, \pi_2, \cdots \pi_N]$ is a probability vector. An agent with multiple prior preference maximizes:

$$\min_{\pi \in \Pi} \left\{ \sum_{i=1}^{N} \pi\left(i\right) u\left(C_i\right) \right\},$$

where Π is a set of probabilities. That is, the multiple prior agent select from a set of probabilities that minimizes the expectation. One interpretation is that the agent is worried that there is a malevolent nature who works against him or her. Note that this setup effectly increases risk aversion. For example, if you choose a constant consumption plan, then nature cannot really harm you, as the expectation under all probabilities are the same. If you choose a risky consumption plan, the agent can always choose a proability measure that has more weight on the bad state of the world.

One can specify the set of probabilities, Π in many ways. Many non-expected utility models (for example, the cumulative prospect theory of Kahneman and Tversky 1992) boils down to choosing a different Π. In this sense, the multiple prior utility model include many non-expected utility model as special cases.

The robust control model of Hansen and Sargent corresponds to choosing Π to be a set of probabilities "close to" a reference probability measure. The closeness is measured by relative entropy.

We first define relative entropy.

Definition 2 *(Relative Entropy)*

Given a reference probability vector $\{\pi(i)\}_{i=n}^{n}$*, the relative entropy of* $\{p(i)\}_{i=n}^{n}$ *with respect to* $\{\pi(i)\}_{i=n}^{n}$ *is defined as:*

$$R(p \parallel \pi) = \sum_{i=1}^{n} p(i) \ln \frac{p(i)}{\pi(i)}.$$

Remark 1 *(Relative Entropy)*

1. *Denote the density (or in general, the Radon-Nikodym derivative) of* p *with respect to* π *as* $\{m(i)\}_{i=n}^{n}$*, that is,* $\frac{p(i)}{\pi(i)} = m(i)$*, then relative entropy can be written as:*

$$R(p \parallel \pi) = \sum_{i=1}^{n} \pi(i) m(i) \ln m(i) = E[m \ln m]. \tag{11}$$

2. *In robust control models, the probability* π *is interpreted as the "reference model". However, the agent is not completely confident about the reference model, and he or she thinks the "true model" should be close to the reference model.*

3. *Relative entropy can be used as a measure of distance between the true model and the reference model. One can show that* $R(p \parallel \pi) \geq 0$ *and "=" holds if and only* $p = \pi$ *(using Jensen"s inequality). That is, the distance between any* p *and* π *is always positive. Hansen and Sargent use the constraint*

$$R(p \parallel \pi) \leq \eta \tag{12}$$

to define a set of probability measures p *that are "close" to* π*.*

To summarize, the robust control model can be viewed as the special case of the Gilboa-Schmeidler model, where Π is defined by relative entropy:

$$\Pi = all \;\; probabilities \;\; p \;\; that \;\; satisfies \;\; R\left(p \parallel \pi\right) \leq \eta.$$

Remark 2 *Note that this equivalence is true in static models. In dynamic models, not neccessarily so. There is a debate between Hansen-Sargent and Epstein and Shneider in terms of the dynamic consistency of these models. I do not think Hansen and Sargent want to interpret their model as a special case of Gilboa Schmeidler that models preferences. I think they want to interpret it as a way of formalizing "doubts" and "Knightian uncertainty", without resorting to Gilboa and Schmeidler axioms.*

A.2 Two Formulations of Robust Control Problems

This section is based on Hansen and Sargent (2001 AER).

Consider an agent wtih concerns for robustness. He calculates his utility level by:

$$\min E\left[mu\left(C\right)\right] \tag{13}$$

$$subject\ to: E\left[m \ln m\right] \leq \eta \tag{14}$$

$$E\left[m\right] = 1 \tag{15}$$

In the above model, I rewrite the relative entropy constraint as a constraint on the density of the probability. This is what Hansen and Sargent (AER 2001) call the *"constraint robust control problem"*, or the *"constraint robust control preference"*. Here η is the parameter that quantify the concern for robustness. Higher η corresponds to more model uncertainty, because it means the set of possible models is large.

The Lagrangian equation for the optimization problem is:

$$\mathcal{L} = E\left[mu\left(C\right)\right] - \mu\left\{E\left[m\right] - 1\right\} + \theta E\left[m \ln m\right].^{2} \tag{16}$$

Using the first order conditions, we have:

$$u\left(C_i\right) - \mu + \theta\left(\ln m_i + 1\right) = 0. \tag{17}$$

[2] Note that I omitted the term η, because this does not affect the optimality conditions.

Together with the condition $E[m] = 1$, (17) imply

$$m_i = \frac{e^{-\frac{u(C)}{\theta}}}{E\left[e^{-\frac{u(C)}{\theta}}\right]}. \tag{18}$$

We use (16) and (18) to write the Lagrangian as:

$$\mathcal{L} = -\theta \ln E\left[e^{-\frac{u(C)}{\theta}}\right]. \tag{19}$$

Of course, to complete the solution to the optimization problem (13-15), we need to express θ as a function of η. However, Hansen and Sargent noted that the funtional form of (19) is much simpler to deal with if we started with θ but not η. So instead of using (13-15) to represent the preference, Hansen and Sargent use the following multiplier robust control problem to represent preferences:

$$\min E[mu(C)] + \theta E[m \ln m] \tag{20}$$

$$\textit{subject to}: E[m] = 1. \tag{21}$$

Problem (20-21) is what Hansen and Sargent call the multiplier robust control problem, or the multiplier robust control preference. As we can see this formulation is purely motivated by mathematical convenience. Hansen and Sargent mostly use the multiplier robust control problem in subsequent work due to its tractability.

Remark 3 *(Multiplier Robust Control Preference)*

1. *Note that the above discussion implies that the solution to the multiplier robust control problem has a closed form:*

$$-\theta \ln E\left[e^{-\frac{u(C)}{\theta}}\right] = \begin{array}{c} \min E[mu(C)] + \theta E[m \ln m] \\ \textit{subject to}: E[m] = 1 \end{array}. \tag{22}$$

   This is the reason why Hansen and Sargent prefer this formulation. Under this formulation, θ is the parameter that quantifies robustness. A large θ corresponds to less model uncertainty. In fact, $-\theta \ln E\left[e^{-\frac{u(C)}{\theta}}\right] \to E[mu(C)]$ (expected utility) as $\theta \to \infty$.

2. *Hansen and Sargent (JET 2007) define an operator T_θ that maps a stochastic utility*

into a certainty equivalent that reflects model uncertainty:

$$T_{\theta}(u) = -\theta \ln E\left[e^{-\frac{u(C)}{\theta}}\right].$$

This is motivated by (22).

3. *Consider a robust control prefrence agent who faces a deterministic consumption C_0 today, and a stochastic consumption C_1 tomorrow. Let's assume his discount rate is β. In this case, the agent calculate his utility according to:*

$$u(C_0) - \theta\beta \ln E\left[e^{-\frac{1}{\theta}u(C_1)}\right]. \tag{23}$$

Let's assume that $u(\cdot)$ is log, and denote $\theta = \frac{1}{\gamma-1}$, for some $\gamma > 1$.[3] *The above can be written as:*

$$\ln(C_0) + \beta\frac{1}{1-\gamma}\ln E\left[C_1^{1-\gamma}\right]. \tag{24}$$

Note that (24) is the utility function of a recursive utility maximizer with unit IES and with risk aversion $\gamma = 1 + \frac{1}{\theta}$. This is the equivalence between robust control of recursive utility.

[3] Note that θ is a Lagrangian multiplier, so $\theta > 1$.