Ambiguity and the historical equity premium

Fabrice Collard
(Toulouse School of Economics, CNRS)

Sujoy Mukerji
(Department of Economics and University College, Oxford U.)

Kevin Sheppard
(Department of Economics, Oxford U.
and Oxford-Man Institute of Quantitative Economics)

Jean-Marc Tallon
(Paris School of Economics, CNRS & U. Paris I)

Preliminary/Work In Progress
Questions addressed in this paper

- How much of the equity premium puzzle may be accounted for by
  1. Uncertainty about the probability distribution over future growth in consumption and dividend
  2. Agents’ sensitivity to this uncertainty
A canonical example of ambiguity aversion

Ellsberg’s 2-color, 2-urn experiment

- A ball is drawn from each urn, generating events $\text{IR}, \text{IB}, \text{II}_R, \text{II}_B$.
- Decision maker is offered bets on these events: £10 if $\text{IR}$, 0 otherwise.
- Typical preferences are: $\text{IR} \succ \text{II}_R$ and $\text{IB} \succ \text{II}_B$
Ellsberg's 2-color, 2-urn experiment

- Modal preferences are: $I_R \succ II_R$ and $I_B \succ II_B$

- Such preferences are not Subjective Expected Utility

- Suppose, we assume $Pr(I_R) = Pr(I_B)$, then the preferences imply

  $$Pr(II_R) < 0.5 \text{ and } Pr(II_B) < 0.5$$

- A single prior cannot express the aspects of uncertainty taken into account by the decision maker.

- The ambiguity averse decision maker takes into account subjective uncertainty about the odds and to what extent the choice is robust to this uncertainty
Motivation for the question

- Pervasiveness of ambiguity aversion well documented in laboratory
- Substantial *theoretical* work on ambiguity aversion
  - foundations and applications to economic models/environments show rich potential
  - theoretical claim that ambiguity aversion could help explain (part of) equity premium but few quantitative papers
- Our paper
  - Use a standard Lucas tree model but for agent preferences and consumption process
  - Quantify ambiguity and ambiguity aversion
  - Special attention to specification that are empirically relevant
  - Provide quantitative assessment under these specifications
  - Analysis does not rely on “tail” events
    * Rare events are interesting but not our story
    * Empirically rare event probabilities cannot be reliably identified
Closely related literature

- Hansen and Sargent (2008):
  - Their question/focus is somewhat broader: not purely ambiguity aversion in the Ellsbergian sense
  - Log-exponential case of KMM coincides with one of their formulations
  - Their operator ($T^1$) seems to be more along the lines of Kreps and Porteus/Epstein and Zin/Weil, not Ellsbergian ambiguity aversion.
Closely related literature

- Hansen and Sargent (2008):
  - Their question/focus is somewhat broader: not purely ambiguity aversion in the Ellsbergian sense
  - Log-exponential case of KMM coincides with one of their formulations
  - Their operator ($T^1$) seems to be more along the lines of Kreps and Porteus/Epstein and Zin/Weil, not Ellsbergian ambiguity aversion.

- Ju and Miao (2008)
  - Also apply the smooth ambiguity framework to connect ambiguity to equity premium
  - We use an empirical specification/stochastic evolution of beliefs that is more compatible with observed consumption that allows for a continuum of states
  - We use a different parametric specification of ambiguity aversion (exponential-power).
• Bansal and Yaron (2004)
  – Introduced “long-run risk” stochastic specification
  – Used to investigate the implications of long run risk rather than state uncertainty
A framework that allows for ambiguity preferences

Smooth ambiguity preferences (KMM (2005)) are represented as,

\[ V(f) = \int_{\theta \in \Theta} \phi \left( \int_{S} u(f(s)) \, d\pi_{\theta}(s) \right) \, d\mu(\theta). \]

- \( s \in S \) set of contingencies or states
- \( \pi_{\theta} \) is a probability distribution over \( S \)
- \( f \) is an “act” yielding state contingent payoffs \( f(s) \)
- \( u \) is a vN-M utility function and represents risk attitude
- \( \phi \) maps expected utilities and represents ambiguity attitude
- Ambiguity attitude is summarized using measure similar to Absolute Risk Aversion, only Absolute Ambiguity Aversion: \( \alpha = -\frac{\phi''}{\phi'} \),
- \( \mu \) is a subjective probability over \( \theta \in \Theta \);
  - Represents the ambiguous belief – it summarizes the DM’s subjective uncertainty about the “true” \( \pi_{\theta} \), probability distribution over contingencies
Recursive Smooth Ambiguity Preferences

- KMM (2008) extend KMM (2005) to intertemporal preferences

- Represented using recursive formulation

\[ V_{s^t}(f) = u(f(s^t)) + \beta \phi^{-1} \left[ \int_{\Theta} \phi \left( \int_{\Pi_{t+1}} V_{(s^t, n_{t+1})}(f) d\pi_{\theta}(n_{t+1}; s^t) \right) d\mu(\theta | s^t) \right] , \]

- \( V_{s^t}(f) \) is a recursively defined (direct) value function
  - \( \beta \) is the discount factor
  - \( \pi_{\theta}(n_{t+1}; s^t) \) is one possible probability distribution at node \( s^t \) on (immediate) successor nodes \( (n_{t+1}; s^t) \)
  - \( \mu(\theta | s^t) \) is the Bayesian posterior given history of observations at node \( s^t \).
  - \( s^t \) observed while \( \theta \) never observed.
Lucas tree

- Representative agent model. Smooth ambiguity preferences.
- Tree that yields uncertain dividend + risk free asset.
- Ambiguity about the stochastic process followed by the tree: need to be more specific about stochastic structure of the economy.
Risks for the Long-run

- Adopt the LRR formulation of Bansal & Yaron (2004) without stochastic volatility
- Consumption and Dividend Growth contain both short term and long term components
- Empirically it is very difficult to distinguish LRR from approx. i.i.d. DGP

Latent, persistent state of the economy

\[ x_{t+1} = \rho x_t + \sigma_x \epsilon_{x,t+1} \]

Consumption growth

\[ g_{t+1} = \bar{g} + x_{t+1} + \sigma_g \epsilon_{g,t+1} \]

Dividend growth

\[ g^d_{t+1} = \bar{g}^d + \psi x_{t+1} + \sigma_d \epsilon_{d,t+1} \]

\[ \epsilon_{x,t+1} \perp \perp \epsilon_{d,t+1}, \epsilon_{g,t+1} \perp \perp \epsilon_{d,t+1}, \epsilon_{g,t+1} \perp \perp \epsilon_{x,t+1} \sim N(0,1) \]
Data

Annual data from 1930 – 2007 (78 observations)

- Asset returns and dividends (CRSP)
- Consumption on nondurables and services (BEA)
- Measures on a per-capita basis using population (BEA)
- Everything measured in real terms using core CPI (no energy or food) (BLS)
Per capita GDP growth
### Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate 1</th>
<th>Estimate 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>.85</td>
<td>.90</td>
</tr>
<tr>
<td>$\psi$</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>$\mu_g$</td>
<td>-0.029</td>
<td>-0.018</td>
</tr>
<tr>
<td>$\mu_d$</td>
<td>-0.138</td>
<td>-0.098</td>
</tr>
<tr>
<td>$\sigma^2_g$</td>
<td>$0.6260 \times 10^{-4}$</td>
<td>$0.6809 \times 10^{-4}$</td>
</tr>
<tr>
<td>$\sigma^2_d$</td>
<td>0.0116</td>
<td>0.0079</td>
</tr>
<tr>
<td>$\sigma^2_x$</td>
<td>$0.87071 \times 10^{-3}$</td>
<td>$0.5450 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

- Parameters estimated using MLE except $\rho$ and $\psi$
- Based on data from 1930 until 1976 – premium and risk free rate using “out-of-sample” data
  - $\rho = .85$ corresponds closely to monthly $\rho = .98$ (Bansal & Yaron (2004))
  - Both $\rho$s lower than H&S’s (2008) quarterly $\rho$ of .98
Smooth ambiguity aversion and long run risk

\[ V_{s^t} (C^t) = u(C_t) + \beta \phi^{-1}\left[ \int_X \phi \left( \int_{\Pi} V_{(s^t, g_{t+1})} (C_{t+1}) \, d\pi_{x_{t+1}} \right) \, d\mu(x_{t+1} \mid s^t) \right], \]

- \( s^t \) represents the history of observed growth “signals”; \( g_{t+1} \) growth rates.
- \( \pi_{x_{t+1}} \) is the distribution of \( g_{t+1} \) given \( x_{t+1} \).
- \( \mu(\cdot \mid s^t) \) is the agent’s beliefs over \( x_{t+1} \).
- Agent never learns true \( x_t \); at node \( s^t \), current belief about the mean of the latent state of economic activity, \( \hat{x}_t \) (enough to specify agent’s beliefs given specification: variance is known, need only to update mean of the distribution)
- \( X \) represents the uncertainty about the current state of the economy through \( x_t \) – without \( \phi \) this would be the usual “Bayesian” adjustment.
Implementation

- CRRA Utility: $u(C_t) = \frac{C_t^{1-\gamma}}{1-\gamma}$

- Constant Absolute Ambiguity Aversion (CAAA): $\phi(u) = -\frac{\exp(-\alpha u)}{\alpha}$

- Bellman formulation: take as “state variables” current wealth and current beliefs $\mu_t$ on $x_t$, summarized by mean $\hat{x}_t$.

- Updating of $\hat{x}_t$ by Kalman filter.
Euler equations

Familiar Euler equations,\(^a\) only with a "twist"

\[
\beta R_t^f E_{\mu_t} \xi_t(x_t | s_t) \left[ E_{\pi x_t} \left( \frac{u'(C_{t+1})}{u'(C_t)} \right) \right] = 1
\]

\[
\beta E_{\mu_t} \xi_t(x_t | s_t) \left[ E_{\pi x_t} \left( R_t \frac{u'(C_{t+1})}{u'(C_t)} \right) \right] = 1
\]

where

\[
\xi_t(x_t | s_t) = \frac{\phi'(E_{\pi x_t} (V(s_t, g_{t+1}) (C_{t+1}))))}{E_{\mu_t} \left[ \phi'(E_{\pi x_t} (V(s_t, g_{t+1}) (C_{t+1})))) \right]}
\]

Note \(V(s_t, g_{t+1}) (C_{t+1})\) is the value at node \((s_t, g_{t+1})\) of the optimal consumption stream which is here identical to the one generated by the tree...

Computational challenge: compute this (direct) value function (recursive, but increasingly difficult as the stochastic specification becomes richer).

\(^a\)Uses the fact that \(\phi\) is exponential. If not, have an extra term in Euler equations.
Intuition behind the "twist"

\[ \xi_t(x_t|s^t) = \frac{\phi'(E_{\pi x_t} (V_{s^t,g_{t+1}} (C^{t+1})))}{E_{\mu t} [\phi'(E_{\pi x_t} (V_{s^t,g_{t+1}} (C^{t+1})))]} \]

- \( \xi_t(x_t|s^t) \) is a change of measure that distorts the Bayesian posterior towards hidden states \( x_t \) with lower continuation values (recall \( \phi' \) decreasing.)

- It operates period-by-period
  - Therefore not classic pessimism
  - Cannot generate these Euler equations from any Bayesian update (except in special cases)
  - Curvature of \( \phi \) around value function at current state determines the degree of distortion
Computing the “direct” value function

- What is the value of holding the tree?
- At time $t$, observe growth rate of consumption and dividend
- Use these observations to update beliefs about $x_t$ (Kalman filter)
- Given our specification, beliefs about $x_t$ follows a normal distribution: can be summarized by one number, its mean (given that the variance is taken to be the “stationary” variance.)

$$V(C_t; \hat{x}_t) = \frac{(1 - \beta) C_t^{1-\gamma}}{1 - \gamma} - \frac{\beta}{\alpha} \ln \left( \int_{-\infty}^{\infty} \exp \right)$$

$$- \left( \alpha \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} V(C_{t+1}; \hat{x}_{t+1}) dF(\epsilon_{g,t+1}) dF(\epsilon_{x,t+1}) dF(\epsilon_{d,t+1}) \right) dF(x_t)$$

$$C_{t+1} = C_t \exp \left( \bar{g} + \rho x_t + \sigma_g \epsilon_{g,t+1} + \sigma_x \epsilon_{x,t+1} \right)$$

The probability density of the latent state, $dF(x_t)$ is the Normal density $N(\hat{x}_t, \hat{P})$. 
Estimating $\hat{x}_t$
The operation of smooth AA
The distorted posterior

- "Rational expectations": $g_{t+1} \sim N(\rho x_t + \bar{g}, \sigma^2_x + \sigma^2_g)$ with $x_t = \hat{x}_t$.

- Bayesian: $g_{t+1} \sim N(\rho \hat{x}_t + \bar{g}, \sigma^2_x + \sigma^2_g) \otimes N(\hat{x}_t, \hat{P})$

- Twisted: $g_{t+1} \sim N(\rho \hat{x}_t + \bar{g}, \sigma^2_x + \sigma^2_g) \otimes N(\hat{x}_t, \hat{P}) \otimes \xi_t(x_t | C^t)$

- Rat. Ex $\rightarrow$ Bayesian (ambiguity): increases variance, mean constant

- Bayesian $\rightarrow$ Twisted (ambiguity aversion): left shift, variance constant
Which values of ambiguity aversion?

- Consider an agent with consumption $W$ who faces a prospect of either 2% or 0% in consumption growth in the following period.

- Suppose this agent is a KMM type with

$$\phi(x) = -\frac{\exp(-\alpha x)}{\alpha} \quad \text{and} \quad u(x) = \frac{x^{(1-\gamma)}}{1-\gamma}$$

- Two scenarios (think back to Urn I and Urn II):
  - 2% with probability 0.5 and 0% with probability 0.5. Call the certainty equivalent for this bet $CE(0)$.
  - 2% with probability $\pi$, where $\pi \in [0, 1]$ with uniform probability. Call the certainty equivalent for this bet $CE(\alpha)$

- Consider the statistic, “Ambiguity premium as a proportion of the expected value of the bet”: $\frac{CE(0) - CE(\alpha)}{.5 \times 2\%} = \frac{A}{100}$

- Camerer (1999) reports experimental data suggest $A$ to be around 10
Calibrating ambiguity aversion

• Calibrated with our representative agent specification at a typical node

• The certainty equivalent of the typical draw (of consumption in the following period):

\[ CE(\alpha) \equiv u^{-1} \left( \phi^{-1} \left( \int \phi \left( \int u(C \exp(g)) \, dF(g; x) \right) \, dF(x) \right) \right) \]

• The certainty equivalent of the same draw assuming all the uncertainty is pure risk:

\[ CE(0) \equiv u^{-1} \left( \int u(C \exp(g)) \, dH(g) \right) \]

where \( H(g) \equiv N(\rho \hat{x} + \bar{g}, \sigma_x^2 + \sigma_g^2) \).

• The statistic is now:

\[ \frac{CE(0) - CE(\alpha)}{E[C|x = \hat{x}]} = \frac{A}{100} \]

where \( F(g; x) \equiv N(\rho x + \bar{g}, \sigma_x^2 + \sigma_g^2) \), \( F(x) \equiv N(\hat{x}, \hat{P}/(1 - \rho^2)) \).
Calibrating risk and ambiguity aversion

Still checking numbers...

Note that $\alpha$ cannot be interpreted without fixing $\gamma$ (i.e., a decision maker characterized by $(\alpha, \gamma)$ cannot be compared in terms of ambiguity aversion to a decision maker characterized by $(\alpha', \gamma')$).

For say $\gamma = 2$, values of $\alpha$ between 1 and 5 give reasonable premium.

Use $\beta = .98$ (not part of the calibration exercise).
Estimating the rates of return

The Euler equations may be rewritten given our stochastic specification as follows:

\[ \beta R_t E_{x_t} \xi_{x_t} (C_t) \left[ E_{x_t} \left( \frac{U'(C_{t+1})}{U'(C_t)} \right) \right] = 1 \] (1)

\[ \beta E_{x_t} \xi_{x_t} (C_t) \left[ E_{x_t} \left( R_t \frac{U'(C_{t+1})}{U'(C_t)} \right) \right] = 1 \] (2)

where

\[ \xi_{x_t} (C_t) = \frac{\phi'(E_{x_t} (V_{t+1}(C_{t+1})))}{E_{x_t} [\phi'(E_{x_t} (V_{t+1}(C_{t+1)))]} \] (3)

- \( V_{t+1}(C_{t+1}) \) is the recursive direct value function which is solved using an exponential specification based on Hermite Polynomials

- **Risk**: \( E_{x_t} \) takes expectations with respect to the probability \( N(\bar{\sigma} + \rho x_t, \sigma^2_g + \sigma^2_x) \)

- **Uncertainty**: \( E_{x_t} \) takes expectations with respect to the probability \( N(\hat{x}_t, \hat{P}) \)
Estimating the rates of return along the sample path

- The first step in computing the prediction on prices is the computation of the direct value function.
- Then, given a sequence \( \{C_t, D_t, \hat{x}_t\}_{t=1}^{t_N} \), compute the associated time \( t \) probability distortion function sequence, \( \{\xi_{\hat{x}_t}(C_t)\}_{t=t_1}^{t_N} \).
- Plugging these in the equations of rates of return and solving, we obtain the time series of moments of rates of return along the observed, actual history.
Overview of results

Use the values of the parameters estimated on period 1930-1976 to predict rates of returns on 1977-2007. Conditional Moments. $\rho = 0.85$.

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>$\alpha$</th>
<th>$r^f$</th>
<th>$r$</th>
<th>$r - r^f$</th>
<th>$r^f$</th>
<th>$r$</th>
<th>$r - r^f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>80</td>
<td>1.0041</td>
<td>4.5234</td>
<td>3.5194</td>
<td>0.4700</td>
<td>1.6187</td>
<td>1.6187</td>
</tr>
<tr>
<td>0.7</td>
<td>42</td>
<td>1.0419</td>
<td>4.6642</td>
<td>3.6222</td>
<td>0.6671</td>
<td>2.2982</td>
<td>2.2982</td>
</tr>
<tr>
<td>0.9</td>
<td>23</td>
<td>1.0935</td>
<td>4.8117</td>
<td>3.7182</td>
<td>0.8921</td>
<td>2.9865</td>
<td>2.9865</td>
</tr>
<tr>
<td>1.5</td>
<td>10</td>
<td>1.3460</td>
<td>5.2395</td>
<td>3.8935</td>
<td>1.5495</td>
<td>5.0355</td>
<td>5.0355</td>
</tr>
<tr>
<td>2.0</td>
<td>4.5</td>
<td>1.0986</td>
<td>5.5381</td>
<td>4.4395</td>
<td>2.1178</td>
<td>6.7628</td>
<td>6.7628</td>
</tr>
<tr>
<td>2.5</td>
<td>1.8</td>
<td>1.0272</td>
<td>5.8996</td>
<td>4.8724</td>
<td>2.6697</td>
<td>8.4983</td>
<td>8.4983</td>
</tr>
<tr>
<td>B&amp;Y Data</td>
<td>.86</td>
<td>7.19</td>
<td>6.33</td>
<td></td>
<td>.97</td>
<td>19.42</td>
<td>19.42</td>
</tr>
</tbody>
</table>

(1929-1998)
Overview of mechanisms

Effects can be understood by simply looking at the distortion picture(s): it is standard risk aversion theory with the twisted distribution (after you made the distortion don’t think about ambiguity aversion anymore, just risk aversion)!

- Given risk aversion, an increase in ambiguity aversion yields:
  - smaller risk free rates: more ambiguity averse agent puts more weight on low continuation value, “saves” more, driving risk free rate downward
  - slightly smaller risky rates
  - increase in the equity premium

- With ambiguity but no ambiguity aversion (i.e., “pure Bayesian case”):
  - averages of risk free rates much higher than with ambiguity aversion.
  - risky rates and volatilities stay roughly the same.

- Increased volatility results from risk aversion acting in conjunction with the increased volatility (of the mixture distribution) compared to the rational expectations case.
Comparison with Ju and Miao

- Their Log-exponential specification does not yield interesting results for the equity premium.
- The power-power specification requires $\gamma < 1$ [for $\gamma > 1$, higher ambiguity aversion leads to less savings and higher risk free rate].
- Have low volatility. Need high $\alpha$ (48).
- Need a rather low discount rate ($\beta = .94$).
Discussion

- In this specification AA allows agent to act as if expected growth will be below what her posterior would indicate.

- Bayesian uncertainty is still important for the risky rate
  - Without AA, the Bayesian explanation cannot generate a realistic premium or risk-free rate – times have been too good.

- Coefficient of ambiguity aversion, $\alpha$, and risk aversion, $\gamma$, are not separable
  - Distortion depends on value function which is determined by utility function.

- Long-run risk is only important in the sense that the current state is persistent over the next period and difficult to pin down.

- Different from B&Y’s use of LRR.
Work in progress
Conclusions

• State uncertainty alone cannot address all of the equity premium puzzle

• Leads to a natural extension of Bayesian framework that allows for conditional pessimism

• Operates using standard tools of finance, only on distorted set of beliefs
  – Distortion based on changes to “central” beliefs, not extremes

• In our specification ambiguity aversion decouples short term rates from expected growth

• Current extensions:
  – Allow for uncertainty over $\rho$
    * i.i.d. vs. LRR (Empirically Questionable)
    * Low persistence vs. High persistence (Empirically Supportable)
    * Creates volatility clustering in returns without having stochastic volatility in consumption
  – Term structure of interest rates: Average effect is to steepen the yield curve, and shows substantial variation over the range of observed $\hat{x}_t$s